THE UNIVERSITY *of* EDINBURGH

The roles of the pluripotency associated *Tex19.1* gene in mouse embryonic and germline development

Judith Reichmann

Thesis presented to the University of Edinburgh for the Degree of Doctor of Philosophy
October 2011

Abstract

Chromosome segregation errors that occur in the developing germline generate aneuploidies which are among the leading causes of embryonic lethality, spontaneous abortions and chromosomal disorders, such as Down's syndrome. Compared to other species, human oocytes appear to be particularly prone to suffer chromosome mis-segregation and the risk of aneuploid pregnancies in humans increases drastically with maternal age. Despite its particular importance for human health, relatively little is known about the basis for the high incidence of aneuploidies in human oocytes and the maternal-age effect. The identification and analysis of molecular pathways that promote genetic and chromosomal stability is important for our understanding of mechanisms that lead to aneuploidy and how it can be prevented.

Here, I examine the role of the pluripotency associated *Tex19.1* gene, in preventing aneuploidy during mouse female germ cell development. I demonstrate that *Tex19.1*$^{-/-}$ females are subfertile when mated with wild type males due to defects in chromosome segregation during meiosis. In contrast to *Tex19.1*$^{-/-}$ male gem cells, synaptonemal complex formation appears to be completed normally in *Tex19.1*$^{-/-}$ females but high levels of aneuploidy are evident during the second meiotic stages of oogenesis. The *Tex19.1*$^{-/-}$ females transmit these aneuploidies to their offspring likely resulting in the observed embryonic death and subfertility.

In addition to its role in the female germline, I investigated the function of Tex19.1 during embryonic development. I found that *Tex19.1*$^{-/-}$ knockout mice are born at a sub-Mendelian frequency and this reduction is exacerbated in diapaused embryos, suggesting that Tex19.1 plays a role during a stage where a pluripotent state is maintained for a prolonged period of time. Furthermore, I identified high levels of aneuploidy accumulating in pluripotent stem cells in the absence of Tex19.1.

Previous work demonstrated that the chromosome segregation defects in *Tex19.1*$^{-/-}$ males correlate with retrotransposon upregulation suggesting that Tex19.1 might be part of a host defence system that protects cells against mutagenic endogenous retroviruses. During this project, I investigated the upregulation of retrotransposon expression in more detail and tested at which stage of the retroviral life cycle inhibition through Tex19.1 might occur. I was able to show that Tex19.1 interacts with the E3 ubiquitin ligase, Ubr2, in 293T and ES cells suggesting a role in proteolysis.

The work carried out during the course of this thesis suggests that Tex19.1 functions at multiple points during the germ line cycle. In particular, the defects in chromosome segregation observed in *Tex19.1*$^{-/-}$ female mice advance our understanding of mechanisms that influence the occurrence of aneuploidy. *Tex19.1*$^{-/-}$ female mice could provide a valuable model system for human aneuploidy and shed light on how our germlines have evolved to maintain chromosomal stability through the generations.

Declaration

I declare that the work presented in this PhD thesis is my own except where otherwise stated. This work has not been submitted for any other degree.

Judith Reichmann Edinburgh, 22nd of October, 2011

Acknowledgements

I am deeply grateful to many people without who I would have never been able to finish my thesis.

I want first to thank my supervisor Dr. Ian Adams for the extraordinary amount of time, help and guidance I received during the last three years and of course for the opportunity to pursue this work in his laboratory. In particular I am indebted to Ian for providing me with an excellent atmosphere for doing research and giving me the chance of learning the wonders of early developmental biology, germ cells – in particular meiosis - and mouse work from an exceptional researcher and teacher as himself. He has supported me in all difficulties and often believed more in me than I did myself. There are just not enough words to thank him.

I owe a thankyou to Dr. Howard Cooke and the past and present members of the Adams' laboratory whose support, technical help and ideas have had a substantial impact on this work. Thanks also to all the past and present members of the Chromosomes and Gene Expression section for making my time at the HGU an enjoyable experience and for their all help and advice and their friendship during the last three years. Many thanks to Alan Hart, Janice Young, Emma Murdoch, David Black, Julie Coles, Malcolm Penman, Gary Waugh, Jacek Mendrychowski and John Jardine for all their work and help in the animal house; to Ian Adams, Diana Best, Sara Heras, Ragnhild Eskeland, Allysson Ross and Shelagh Boyle for teaching me lab techniques; to Paul Perry and Matthew Pearson for their master classes in microscopy and all the members of the HGU core scientific services who made my (work) life so much easier. Special thanks to Diana Best, Shelagh Boyle, Marie Maclennan, James Crichton, Simona Paro, Diana Reinhardt, Nele Hug, Sara Heras and Sara Macias for their humour, smile and support. They all made lab-and office work joyous every single day.

I've been very lucky to have some marvelous friends who made life outside the lab an unforgettable enjoyable experience. I thank them for their friendship and support over the past years, especially Emma Hall and Anikki Mair who made me feel at home in Edinburgh, and wish them all the best in their own studies.

Finally, thanks go to my husband James Reddington, my sister Christine Reichmann and my parents Ulrich and Cornelia Reichmann. This thesis would never have happened without them. Their unconditional love and support and their unquestioning belief in me mean the world to me. This is for them.

Abbreviations

A	:	Adenine
aa	:	Amino acid
Ab	:	Abembryonic
AE	:	Axial element
AVE	:	Anterior visceral endoderm
Aza	:	5'-deoxy-2'-azacytidine
BLAST	:	Basic Local Alignment Search Tool
bp	:	Base-pair
BSA	:	Bovine serum Albumin
cDNA	:	Complementary Deoxyribonucleic acid
C	:	Cytosine
CE	:	Central element
°C	:	Grad Celsius
ChIP	:	Chromatin-immuno-precipitation
CR	:	Crossover
DMEM	:	Dulbecco's Modified Eagle Medium
dpc	:	Days post coitum
dpp	:	Days post partum
DNA	:	Deoxyribonucleic acid
DSB	:	Double strand break
DTT	:	Dithiothreitol
DVE	:	Distal visceral endoderm
FISH	:	Fluorescence in situ hybridisation
EB	:	Embryoid body
EC	:	Embryonic carcinoma
EDTA	:	Ethylenediaminetetraacetic acid
EG	:	Embryonic germ
Em	:	Embryonic
ERV	:	Endogenous retroviruse
ES	:	Embryonic stem
EtBr	:	Ethidium Bromide
ExE	:	Extra embryonic ectoderm
FCS	:	Fetal calf serum
Fsh	:	Follicle-stimulating hormone
g	:	Gram
GnRH	:	Gonadotropin-releasing hormone
G	:	Guanine
GVBD	:	Germinal vesicle break down
h	:	Hour
hCG	:	Human chorionic gonadotrophin
HDAC	:	Histone deacetylase
HEPES	:	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
H3K9me^2	:	Histone 3 lysine 9 dimethylation
H3K9me^3	:	Histone 3 lysine 9 trimethylation
H3K27me^3	:	Histone 3 lysine 27 trimethylation
i.e.	:	Id est
IHC	:	Immuno-histo-chemistry
ICM	:	Inner cell mass
IF	:	Immuno-fluorescence
IVF	:	*In vitro* fertilisation

kDa	:	Kilodalton
IRES	:	Internal ribosome entry site
L	:	Liter
LE	:	Lateral elements
Lh	:	Leutinizing hormone
LIF	:	Leukemia inhibitory factor
LINE	:	Long interspersed elements
LTR	:	Long terminal repeats
MEF	:	Mouse embryonic fibroblast
$MgCl_2$	:	Magnesium Chloride
MI	:	Meiosis I
MII	:	Meiosis II
µg	:	Micro gram
mL	:	Milli liter
min	:	Minute
mRNA	:	Messenger ribonucleic acid
N	:	Any DNA base (adenine/guanine/cytosine/thymine)
NCBI	:	National center for biotechnology information
NCR	:	Non-crossover
NEAA	:	Non essential amino acids
ORF1	:	Open reading frame 1
ParE	:	Parietal endoderm
pMEF	:	Primary mouse embryonic fibroblast
PAGE	:	Polyacrylamide gel electrophoresis
PBS	:	Phosphate buffered saline
PBST	:	Phosphate buffered saline with tween
PCR	:	Polymerase chain reaction
PGCs	:	Primoridial germ cells
PMS	:	Pregnant mare serum
RNA	:	Ribonucleic acid
RNAi	:	RNA interference
rpm	:	Rounds per minute
SAC	:	Spindle-assembly checkpoint
SC	:	Synaptonemal complex
SDS	:	Sodium dodecyl sulfate
SSC	:	Spermatogonial stem cell
SINE	:	Short interspersed element
T	:	Thymine
TBE	:	Tris/Borate/EDTA
TE	:	Trophectoderm
TEMED	:	Tetramethylethylenediamine
TF	:	Transverse filaments
TKO	:	Triple knockout
TSA	:	Trichostatin A
Tex19.1	:	Testis expressed gene 19.1
UTR	:	Untranslated region
UV	:	Ultra violet
VE	:	Visceral endoderm
XRV	:	Exogenous retrovirus

Table of contents

Chapter 1: Introduction

1.1 The germline cycle

Reproduction is a fundamental feature of all known life. Sexual reproduction is a biological process by which organisms create descendants that have a combination of genetic material contributed from two different members of the species. Germ cells are the founder cells of all sexually reproducing organisms. Specifically, in mammals genetic information is transmitted through successive generations by being passed in a self-perpetuating cycle from germ cells to pluripotent cells in the early embryo and back again (figure1.1). Pluripotent cells are capable of giving rise to any cell type of the adult organism, including germ cells. Germ cells give rise to eggs in females and sperm in males, the cells responsible for passing genetic information from one generation to the next. Fertilisation of the egg by the sperm creates a totipotent zygote which is able to initiate repetition of the whole cycle again. When the genome is passed on to subsequent generations during this cycle both the germ cells, and the pluripotent cells, must retain a high developmental and infinite proliferative capacity, and ensure genetic and chromosomal stability is maintained.

During my PhD I have particularly focused on the question how chromosomal stability is maintained through successive generations by addressing the role of the pluripotency associated *Tex19.1* gene during the germline cycle, in particular during embryonic and female germline development. This chapter reviews the current knowledge of germ cells and early embryonic development as well as the causes and consequences for the occurrence of chromosomal abnormalities during development.

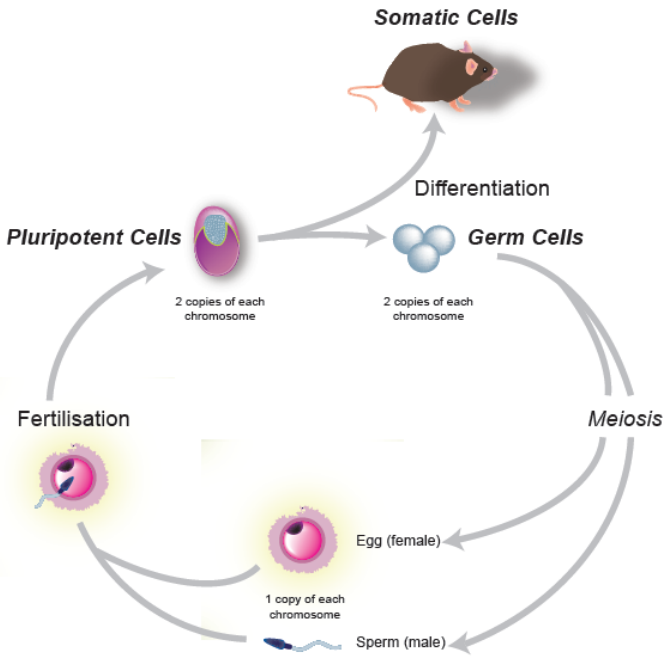


Figure 1.1 The mammalian germline cycle. Fertilization of the egg (pink) by the sperm (blue) produces a pluripotent zygote. The pluripotent cells in the early embryo (blue) will in turn give rise to germ cells (blue). The germ cells will undergo meiosis to produce egg or sperm which will then upon fertilization give rise to the next generation. Modified from (Öllinger et al., 2010).

1.2 Mouse embryonic development

1.2.1 Early mouse pre-implantation development

Embryonic development of the mouse begins with fertilisation of the egg (or oocyte) by the sperm, both of which are haploid cells that no longer divide, resulting in the formation of the zygote (one-cell embryo), a diploid totipotent cell (Seydoux and Braun, 2006). This cell is able to produce all of the tissues necessary for fetal development by having the ability to differentiate into any cell of the organism as well as the extraembryonic tissue associated with the fetus. After fertilisation, the zygote divides into a number of smaller cells called blastomeres, a process which is referred to as cleavage. After fertilisation, the mitotic cell cycle and embryonic developmental

program is resumed which is facilitated by maternally inherited RNAs and proteins (Latham et al., 1992; Latham and Schultz, 2001; Stitzel and Seydoux, 2007). In the mouse the transition from maternal to zygotic control of gene expression occurs at the two-cell stage and during this time the main wave of transcriptional activation and translation of the embryonic genes is initiated (Schultz, 2002). This process, which is also called zygotic genome activation, progressively frees the embryo from the need to use maternally stored ribonucleic acids (RNAs) (Flach et al., 1982; Nothias et al., 1995; Hamatani et al., 2004a). By 2.5 days post coitum (dpc) repeated cleavage has generated eight morphologically discrete blastomeres (shown in figure 1.2). Before the blastomeres divide for the fourth time surface contact between cells increases through increasing cell adhesion and individual blastomeres become indistinguishable (Hyafil et al. 1980; Shirayoshi et al. 1983; Vestweber, et al., 1987). This process is referred to as compaction and induces a degree of cell polarization, as the initial radially symmetric blastomeres become polarised in an apical–basal manner during this time (Vestweber et al., 1987; Fleming et al., 2001). As cleavage proceeds to the 16-cell stage that occurs around 3dpc, there is a restriction in the developmental potency of the blastomeres eventually resulting in the generation of two distinct lineages: the trophectoderm (TE) and the inner cell mass (ICM) (Wang and Dey, 2006).

The TE will go on and form one of the four extraembryonic membranes, the chorion, which provides the embryonic portion of the placenta. The cells of the ICM will give rise to all the cells of the organism plus the other three extraembryonic membranes (yolk sac, amnion and allantois). The ICM cells are considered to be pluripotent because they have the potential to differentiate into all three germ layers of the embryo proper: ectoderm, mesoderm and definitive endoderm. Formation of TE and ICM facilitates embryonic-abembryonic (Em-Ab) polarity in the blastocyst. The Em pole refers to the side of the blastocyst where the ICM is located whereas the Ab pole marks the opposite side where the blastocoel is situated (figure 1.2). The exact mechanisms that underlie lineage specification of TE and ICM, as well as formation of Em-Ab polarity, in cleavage-stage embryos are still not fully understood (Johnson, 2009). However, there

are several models for this process that have been postulated during the last decades that aim to explain what is known to date (Sasaki, 2010).

The earliest proposed model is the mosaic model which suggested that cell lineages are specified by selective distribution of determinants (Dalcq, 1957; Mulnard 1992). This model appears to be true in a variety of non-mammalian species where early cleavages play an important role in establishing Em-Ab polarity. For example in frogs and fish determinants such as β-catenin define formation of the dorsal axis. Removal or ectopic expression of dorsal determinants in frogs or fish results in ventralised or dorsalised embryos, respectively (Pandur et al., 2002; Schier and Talbot, 2005; White and Heasman, 2008). In contrast, in mice, removal of the animal or vegetal pole of the zygote has not shown to negatively affect embryonic development as there is no adverse phenotype, therefore demonstrating that critical determinants are not needed in mouse zygotes, and arguing against the mosaic model for embryonic patterning in mice (Zernicka-Goetz 1998). Alternatively, it has been proposed that the pattern of early cleavages influences lineage specification during mouse embryonic development. Several studies suggest that the plane of the first cleavage event determines that one blastomere of the two-cell embryo will mainly contribute to the ICM and the other mainly to the TE cell lineages (Gardner, 1997; Gardner, 2001; Piotrowska et al. 2001; Plusa et al. 2005; Zernicka-Goetz, 2005). Furthermore, the second cleavage has also been implicated in deciding the developmental fate of the blastomeres in that the cell which divides first will mainly contribute to the ICM whereas the later dividing cell will give rise mainly to the TE (Piotrowska et al., 2001; Piotrowska-Nitsche and Zernicka-Goetz, 2005a; Piotrowska-Nitsche and Zernicka-Goetz, 2005a; Bischoff et al., 2008). This model remains controversial as other studies were not able to identify a correlation between early cleavage and blastocyst organization or a specific fate for early or later dividing blastomeres (Alarcon and Marikawa, 2003; Fujimori et al., 2003; Chroscicka et al., 2004; Alarcon and Marikawa, 2005; Motosugi et al. 2005; Kurotaki et al. 2007). Furthermore, it has been shown that artificial removal of one blastomere at the 2 or 4-cell stage does not impair embryonic development (Tarkowski, 1959). Similarly,

aggregation experiments of single blastomeres, isolated at the 4- or 8-cell stage, with host embryos have shown that those cells are still totipotent suggesting that initially during cleavage all blastomeres are identical in their developmental potential and that linage specification in the mouse occurs after the 8-cell stage (Tarkowski & Wroblewska 1967; Rossant, 1976; Kelly, 1977).

1.2.2 Lineage specification during mouse pre-implantation development

It has long been proposed that blastomeres after the forth cleavage adopt TE or ICM fates depending on their position within the embryo (Tarkowski and Wroblewska, 1967). This model was later refined with the polarisation model (Johnson and Ziomek, 1981a; Johnson and Ziomek, 1981b). The polarisation model proposes that polarisation of the 8-cell stage blastomeres is critical for later lineage determination. While cells divide from 8- to 16-cell stage some blastomeres will acquire a position on the surface of the embryo whereas others will adopt an internal position where they are completely surrounded by neighboring blastomeres (figure 1.2) (Johnson, 2009). The outer blastomeres retain an apical-basal polarity whereas internal blastomeres lose apical features and become morphologically apolar (Johnson and Ziomek, 1983). During the fifth cleavage event more external and internal cells are generated by division of the outer polarised cells. Symmetric division of external cells results in two polarised cells which will stay on the outside. In contrast, asymmetric division of polarised cells gives rise to one polarised daughter which will remain external and one apolar daughter which will adopt an internal position (Johnson and Ziomek, 1981b). External blastomeres preferentially differentiate into TE whereas internal cells go on to form the ICM (figure 1.2) (Johnson and Ziomek, 1983; Johnson and McConnell, 2004; Suwinska et al. 2008). However, until blastocyst formation the fate of individual blastomeres can be changed by changing their position within the embryo (Hillman et al., 1972). During the fifth cleavage event small cavities start to form, between blastomeres, which continually expand and

ultimately fuse to form a single large cavity (blastocoel) (Aziz and Alexandre, 1991). From this point onwards the embryo is termed a blastocyst (figure 1.2).

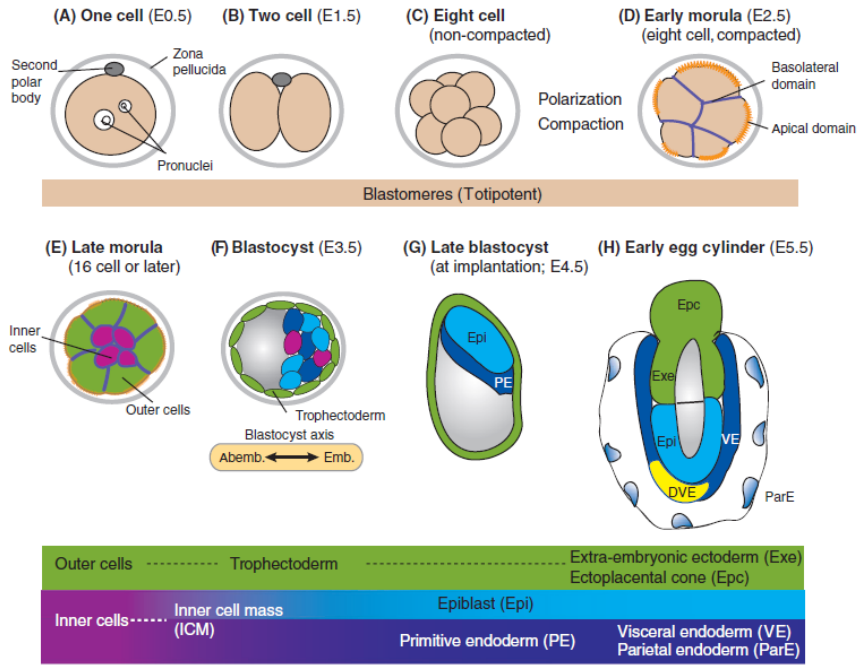


Figure 1.2 Cell lineage formation during mouse embryonic development. Embryonic development of the mouse begins with fertilization of the oocyte by the sperm in the oviduct resulting in a totipotent zygote. The embryo then undergoes several rounds of cell division ultimately resulting in a ball of cells called a morula. The late morula enters the uterus and develops into a blastocyst. The blastocyst contains a cavity (blastocoel) with two distinct populations, the inner cell mass (ICM) and the trophectoderm (the progenitor of trophoblast cells). Around 4 dpc the blastocyst hatches from its outer shell (the zona pellucida) and differentiates to produce additional cell types — the epiblast and the primitive endoderm. The trophectoderm then attaches to the uterine wall to facilitate the process of implantation.The cell types in the embryos are color coded. Figure was taken from Rossant and Tam (2009).

At the blastocyst stage TE cells surround the entire embryo and ICM cells accumulate as a single mass attached to the basal surface of the TE. Lineage specification of TE is initiated at the morula stage prior to blastocyst formation when TE-specific transcription factors like such as *Cdx2* or *Eomes* start to become restricted to the outside blastomeres (Sasaki, 2010). Expression of *Cdx2*, which is detected for the first time at the 8- to 16-cell stage, initially appears random (Ralston and Rossant, 2008). However, from the early morula stage onwards *Cdx2* expression becomes gradually restricted to the outer cells of the embryo. In the outer blastomeres *Cdx2* facilitates suppression of ICM

identity resulting in the restriction of *Oct4* and *Nanog* expression, which are required to specify the pluripotent cells of the ICM, specifically to an ICM fate after blastocyst formation. *Oct4* is expressed throughout early cleavage in all blastomeres as a result of a maternally inherited transcript (Rosner et al., 1990; Schöler et al., 1990; Yeom et al., 1996). At the blastocyst stage *Oct4* becomes restricted to the ICM and is downregulated in the TE cells.

Nanog expression is observed for the first time at the 8-cell stage and once again with an apparently random distribution (Dietrich and Hiiragi, 2007). At the blastocyst stage, similarly to *Oct4*, *Nanog* becomes downregulated in the TE. *Nanog* also becomes downregulated (or fluctuates) in a subpopulation of the ICM (see below). In *Cdx2*$^{-/-}$ mutant mice *Oct4* and *Nanog* do not become downregulated in the TE cells (Ralston and Rossant, 2008). *Cdx2*$^{-/-}$ embryos are able to develop into early blastocysts and form the cavity but the blastocoel eventually collapses as *Cdx2*$^{-/-}$ TE cells fail to undergo further trophoblast differentiation and loose morphological integrity (Strumpf et al. 2005). In accordance with the hypothesis that *Cdx2* is involved in TE lineage specification is data from embryonic stem (ES) cells that suggest that ectopic expression of *Cdx2* induces trophoblast differentiation (Niwa et al., 2005). *Tead4*, *Eomes* and *Elf5* have also been implicated to govern the specification of the TE lineage (Niwa et al., 2000; Niwa et al., 2005). The phenotype of *Tead4*$^{-/-}$ embryos is slightly more severe than *Cdx2*$^{-/-}$ embryos and is characterised by the absence of TE or trophoblast-specific gene expression, failure to form the blastocoel and differentiation of all blastomeres into ICM (Yagi et al. 2007; Nishioka et al. 2008). The apparent roles of *Eomes* and *Elf5* expression are to stabilise the TE lineage and to be required for the differentiation of fate-specified TE into trophoblast cells (Russ et al., 2000; Ng et al., 2008). As described above *Cdx2*$^{-/-}$ mutant embryos fail to silence *Oct4* and *Nanog* expression in the TE lineage. Similarly, *Oct4* and *Nanog* are required to repress *Cdx2* in the pluripotent cells (Niwa et al. 2005; Chen et al. 2009). Failure to repress the expression of *Cdx2* in the inner cell mass, due to mutations in *Oct4,* causes all cells to commit to the TE lineage (see below) (Nichols et

al., 1998). Taken together this data suggests that *Oct4*, *Nanog* and *Cdx2* play crucial roles in segregating pluripotent and TE cell fates, respectively.

1.2.3 Primitive endoderm and epiblast formation

At approximately 4.5dpc the blastocyst is ready for implantation. At this stage the ICM of the blastocyst segregates into epiblast and primitive endoderm (or hypoblast). The epiblast will give rise to all cells of the fetus and extraembryonic mesoderm. The hypoblast will eventually form extraembryonic endoderm layers of the visceral and parietal yolk sacs. *Oct4* has been shown to be absolutely required for the ability of the ICM to differentiate into epiblast and possibly the hypoblast (Nichols et al., 1998). At the late blastocyst stage, when *Oct4* expression is down regulated in the TE lineage, *Oct4* is expressed in all cells of the epiblast and also briefly in the hypoblast. After implantation *Oct4* expression will be restricted to the pluripotent epiblast and is absent from extraembryonic cell lineages (Palmieri et al., 1994). After gastrulation, *Oct4* expression becomes restricted to primordial germ cells (PGCs), the precursors of the germ cell lineage (Schöler et al., 1990; Yeom et al., 1996). As mentioned previously under 1.2.2, $Oct4^{-/-}$ embryos fail to produce the epiblast and hypoblast and instead cells of the ICM differentiate along the extraembryonic trophoblast lineage (Nichols et al., 1998). This suggests that *Oct4* is absolutely required to maintain pluripotency (developmental potential) of the ICM cells in the blastocyst during embryonic development. Consistent with this, ES cells, which are ICM-derived cell lines require a critical amount of *Oct4* expression to maintain pluripotency and expression of *Oct4* is downregulated upon ES cell differentiation (Niwa et al., 2000). Absence of *Oct4* causes ES cells, similarly to ICM cells, to differentiate into TE (Niwa et al., 2000). Based on this data and the transient expression in the hypoblast it has been hypothised that *Oct4* is essential for the differentiation of ICM into epiblast and possibly also for hypoblast commitment (Nichols and Smith, 2009).

Similarly to *Oct4*, *Nanog* is expressed in pluripotent embryo cells, ES cells, and the developing germline, and its presence is associated with an undifferentiated cell state (Chambers et al., 2003, Mitsui et al., 2003). During embryonic development *Nanog*$^{-/-}$ cells initially form pluripotent cells but are unable to maintain this state and fail to form the epiblast (Mitsui et al., 2003). In contrast to *Oct4*$^{-/-}$ ICM cells *Nanog*$^{-/-}$ ICM cells do not differentiate into TE suggesting different functions for these two genes. ES cells derived from *Nanog*$^{-/-}$ blastocysts can be maintained on feeder cells for at least two months but grow slower than control ES cells, down regulate pluripotency markets like *Oct4* and induce expression of endoderm transcription factors and markers (Mitsui et al., 2003). When maintained on gelatin, *Nanog*$^{-/-}$ ES cells completely down regulate expression of pluripotency markers and differentiate into extraembryonic endoderm lineages (Mitsui et al., 2003). This data appears to be in disagreement with a later study which reported that ES cells that have undergone conditional deletion of *Nanog,* after ES cell derivation, are prone to differentiate but are able to proliferate infinitely and contribute to somatic chimaeras (Chambers et al., 2007). However, it has been suggested that *Nanog* functions later than *Oct4* (and other pluripotency factors), which are believed to trigger a cascade of events necessary to induce pluripotency, and instead is required to lock cells into a self-sustaining pluripotent state (Yates and Chambers, 2005; Silva et al., 2008; Nichols and Smith, 2009). Upon differentiation of the epiblast, during gastrulation (see below), *Oct4* and *Nanog* expression is downregulated in all somatic lineages but persists in the primordial germ cells (PGCs) (Chambers et al., 2003; Mitsui et al., 2003; Yamaguchi et al., 2005; Nichols and Smith, 2009).

Morphologically the epiblast and the hypoblast can be clearly distinguished from each other by 4.5dpc. The epiblast and hypoblast appear to be restricted to their corresponding cell fate by 4.5dpc as experiments showed that epiblast or hypoblast cells were able to only contribute to their own lineage when injected into other blastocysts (Moody, 1999). Traditionally it was believed that all cells of the ICM have equal potential to acquire either epiblast or hypoblast cell fate and that the position of cells on the surface of the ICM would be designated to differentiate into hypoblast (Gardner,

1983). Recent studies showed that by 3.5dpc individual ICM cells express mutually exclusive epiblast (*Nanog*) or hypoblast specific genes (*Gata6*) in a random manner (Chazaud et al., 2006). Furthermore, injection of single ICM cells into blastocysts demonstrated that individual cells largely contribute to either hypoblast or epiblast cell lineages (Chazaud et al., 2006). Based on this data it has been postulated that by 3.5pc ICM cells are largely designated to either an epiblast or a hypoblast cell fate and that later relocation of cells segregates the two layers (Chazaud et al., 2006). Some cells in this experiment contributed to both lineages suggesting that by 3.5dpc ICM cells are possibly biased but not absolutely committed to differentiate into either hypoblast or epiblast. This has also been hypothesised by a later study involving lineage tracing of individual ICM cells expressing a marker of the late hypoblast cell lineage (*Pdgfrα*). Live imaging of embryos expressing a histone H2B-GFP fusion protein reporter under the control of the *Pdgfrα* promoter allowed lineage tracing of individual ICM cells (Plusa et al., 2008). This analysis showed that before formation of the blastocyst hypoblast and epiblast specific transcription factors are expressed in an overlapping manner but confirmed mutually exclusive expression by 3.5dpc and that this biases cells towards one or the other lineage. The authors further suggest that at least some aspects of hypoblast induction require positional signals, linking the two hypotheses (Plusa et al., 2008). Further work will be required to understand the exact relationship between lineage restriction and gene expression and the mechanisms involved in segregating cells to their respective layer.

1.2.4 Early post-implantation development

By whichever mechanisms epiblast and hypoblast are segregated it is clear that at implantation the mouse blastocyst is distinctively partitioned into TE, hypoblast and epiblast (figure 1.2). At approximately 4.5dpc the blastocyst is ready for implantation and attaches tightly to the uterine epithelium (Senner and Hemberger, 2010). The TE cells begin to differentiate into an inner layer of cytotrophoblast and an outer multinucleated cell layer, the syncytiotrophoblast which starts to invade the connective

tissue of the uterus (Senner and Hemberger, 2010). Furthermore, subsequent to implantation the polar TE, which lines the epiblast, facilitates the formation of the ectoplacental cone which connects the egg cylinder (an elongated structure consisting of the epiblast, visceral endoderm and the extra-embryonic ectoderm) to the uterine tissue. The TE will ultimately give rise to the trophoblast lineages that form the majority of the fetal part of the placenta. The hypoblast gives rise to the visceral and parietal endoderm (VE and ParE) which surround the epiblast and the extraembryonic ectoderm and line the luminal surface of the mural trophectoderm, respectively (Rossant and Tam, 2009).

The majority of the VE will give rise to extra-embryonic cell lineages, a small proportion of VE cells, however, will also contribute to the endoderm of the embryonic gut (Kwon et al., 2008; Rossant and Tam, 2009). The VE is an important source of signals to establish anterior-posterior polarity of the mouse embryo which is required for formation and correct placement of the primitive streak which marks the initiation of gastrulation (Tam and Loebel, 2007). The primitive streak forms opposite the anterior visceral endoderm (AVE) which defines the anterior side of the embryo (Thomas and Beddington, 1996; Rivera-Perez et al., 2003; Srinivas et al., 2004; Torres-Padilla et al., 2007). The AVE arises from visceral endoderm cells that form at the distal tip of the embryo (also called distal visceral endoderm (DVE)) and then migrate to a more proximal position close to the extra embryonic ectoderm (ExE). Signals from the ExE, like *Bmp4*, induce expression of primitive streak markers like *Wnt3* and *Noda*l in the epiblast. In turn *Nodal* signaling from the epiblast maintains *Bmp4* expression in the ExE (Liu et al., 1999; Ben-Haim et al., 2006). Secretion of Nodal and Wnt3 antagonists like Cer1, Lefty1 and Dkk1 by the AVE confers anterior identity to the underlying epiblast by shielding it from signals that induce primitive streak formation in the posterior epiblast and restrict the primitive streak to the appropriate location within the embryo (Yamamoto et al., 2004; Kimura-Yoshida et al., 2005; Rodriguez et al., 2005; Richardson et al., 2006). Indeed, in embryos that lack both antagonists, the anterior side of the primitive streak becomes expanded, or ectopic primitive streaks form (Perea-Gomez et al., 2002). The primitive streak is the hallmark for the initiation of gastrulation

which results in the formation of the embryonic germ layers (ectoderm, mesoderm and definitive endoderm) (Tam and Loebel, 2007). Epiblast cells that ingress through the primitive streak are incorporated into the endoderm or form a new layer of mesoderm in the space between the epiblast and the definitive endoderm (Tam and Loebel, 2007). Cells that remain in the epiblast will form the ectoderm. Lineage tracing of individual cells showed that epiblast cells are able to give rise to any of the three germ layers until ingression takes place (Lawson et al., 1991). A diagram summarising the process of early post-implantation development in the mouse is shown in figure 1.3. In addition to the three germ layers are primordial germ cells (PGCs), which constitute the founder cells of the germ cell lineage, established during gastrulation around 7dpc (Tam and Loebel, 2007). This process, also referred to as germ cell specification, is described in detail later under 1.3.

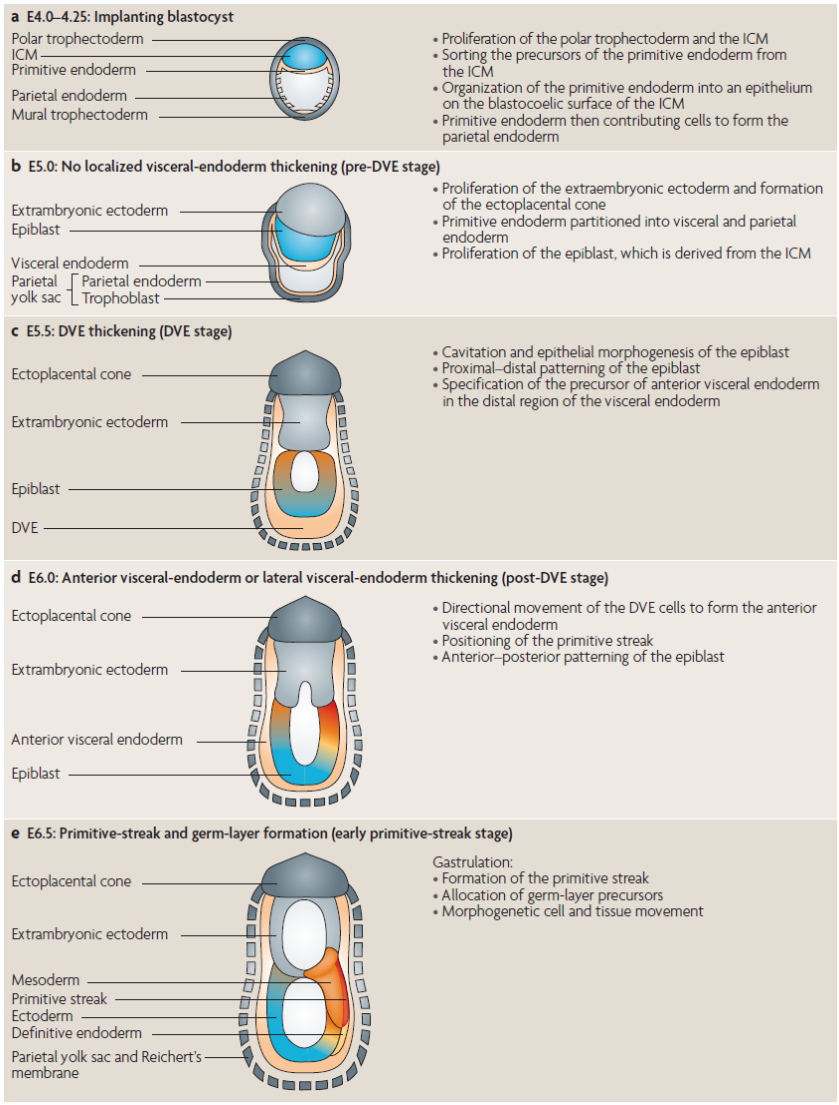


Figure 1.3 Early post-implantation development from 4dpc to 6.5dpc. (a) Blastocyst at implantation (see also figure 1.2). (b) 5dpc embryo prior to formation of distal visceral endoderm (Pre-DVE). (c) The 5.5dpc embryo is characterised by formation of the distal visceral endoderm (DVE). (d) At 6dpc the anterior visceral endoderm has formed (AVE). (e) Early primitive streak formation at 6.5dpc. Primitive endoderm is presented in grey. In a-b ICM and epiblast in blue. In c-e ectoderm, mesoderm and definitive endoderm progenitors are marked in blue, orange and yellow respectively. Primitive streak is colored in crimson. Figure was taken from Tam and Loebel (2007).

1.2.5 Diapause

The above sections described stages of embryonic development that normally occur sequentially and without interruption. However, rodents have the ability to delay embryonic development by arresting the embryo at the late blastocyst stage, after segregation of epiblast and hypoblast but prior to implantation to counteract sub-optimal conditions for reproduction associated with demands on maternal nutrients (Renfree and Shaw, 2000). For example in mice diapause occurs to delay implantation of newly formed embryos in response to lactation. This might occur frequently when a mating pair is housed together as female mice are able to mate shortly after the female has given birth during the post-partum estrus. Implantation itself requires a synchronous interplay between embryos and the uterus as the uterus is only capable of supporting attachment and invasion by the blastocyst for a limited period of time. In mice, uterine receptivity to accept a blastocyst for implantation depends on the coordinated actions of the ovarian steroid hormones, progesterone and estrogen, as implantation fails in estrogen and progesterone mutant mice (figure 4.8) (Lyden et al., 1995; Hewitt et al., 2002). Preparation for implantation begins with a preovulatory increase in estrogen which initiates proliferation of uterine epithelial cells. By 3 dpc, progesterone levels begin to increase which stimulates uterine stromal cell proliferation. High levels of progesterone are needed for both embryo implantation and maintenance of post-implantation development. By 4dpc, estrogen levels increase again which appears to be critical for implantation to occur and stimulate, together with progesterone, the proliferation and differentiation of uterine stromal cells. This rise in estrogen concentration does not take place when the mating has occurred during post partum estrus. This is due to the fact that lactation inhibits the secretion of gonadotropin, which in turn stimulates the production of estrogen, resulting in insufficient secretion of ovarian estrogen at 4 dpc in mothers that are suckling a litter (figure 4.8). As soon as estrogen levels increase, implantation proceeds normally. Diapause can be induced experimentally by ovariectomy of female mice at 2.5dpc, which abolishes the rise in estrus at 4 dpc, and simultaneous administration of progesterone to maintain the pregnancy. Implantation

and normal development of experimentally arrested embryos can be achieved by pre-treatment with progesterone followed by small dose of estrogen or by transfer to a hormonally primed, pseudopregnant recipient mother (Yoshinaga & Adams, 1966; McLaren, 1973). Diapause can also be induced by chemical ovariectomy by simultaneous administration of tamoxifen and depo-provera on day 2.5 of gestation (Hunter and Evans, 1999). Tamoxifen is a non-steroidal anti-estrogen that has been shown to exhibit estrogen antagonistic and agonistic actions in a species specific manner (Hunter and Evans, 1999). Combined administration of tamoxifen and depo-provera (the brand name for depot medroxyprogesterone acetate), appears to cause anti-estrogenic effects resulting in delayed implantation (Hunter and Evans, 1999). Similarly to embryos diapaused by surgical ovariectomy chemically diapaused embryos implant and develop to term when transferred into pseudopregnant recipient mother at the right stage of gestation.

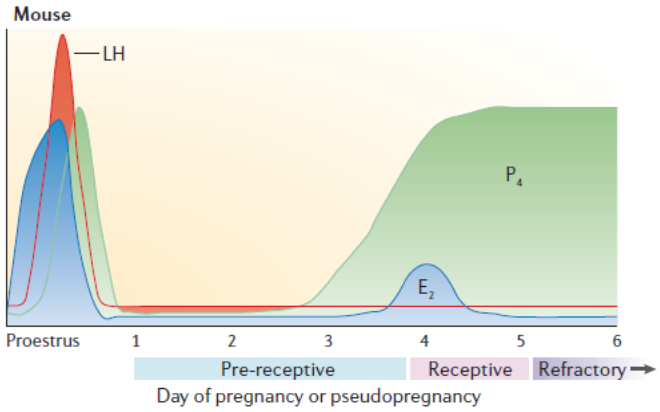


Figure 1.4 Hormone levels required for embryonic development and implantation. A rise in leutinizing hormone (Lh) levels, a gonadotropin (orange), stimulates estrogen (blue) and progesterone (green) secretion by the ovary resulting in ovulation. Preparation for implantation begins with the preovulatory increase in estrogen which initiates proliferation of uterine epithelial cells. By 3 days post-coitum progesterone levels begin to increase (P_4) which stimulates uterine stromal cell proliferation. During this time the uterine environment is considered pre-receptive which means that the uterus is unable to initiate implantation but also less hostile to blastocyst survival. By 4 dpc of pregnancy, or pseudopregnancy, estrogen levels increase again (E_2) and stimulate together with progesterone the proliferation and differentiation of uterine stromal cells resulting in a uterine environment that is receptive for blastocyst implantation. By 5dpc the uterine environment becomes nonreceptive (refractory) to implantation. Lactationally, surgically and chemically-induced diapauses antagonize the second peak in estrogen levels (E_2) and inhibit blastocyst implantation. Simultaneous administration of progesterone ensures that the pregnancy is maintained; figure taken from (Wang and Dey, 2006).

Entry of the embryo itself into diapause may involve the epiblast cells entering or perpetuating an ES cell-like state (Batlle-Morera et al., 2008). This is based on the observation that the LIF/gp130 signalling pathway, which is important for the maintenance of ES cells in culture, is required for diapause, but not for normal development at this stage. The main function of gp130 signalling in vitro is the suppression of ES cell differentiation (Smith, 2001). It is possible that the phenotype of *gp130*$^{-/-}$ embryos is caused by apoptosis of inappropriately differentiating epiblast cells (Nichols et al., 2001). Apart from the gp130 receptor little is known about the genes that the embryo requires to undergo diapause. It is known that upon diapause the blastocyst enters a state of proliferative and metabolic quiescence. Expression changes of many genes have been observed between activated and delayed blastocysts by microarray (Hamatani et al., 2004). However, many of those expression changes can be attributed to the cellular and physiological events that are expected to change with ceased proliferation and metabolism. Of the 229 genes that were found to be differentially expressed between delayed and activated blastocysts approximately 30% have not been functionally investigated so far. Therefore it is possible that one or several important regulators of diapause might be among those (Hamatani et al., 2004). Interestingly, expression levels of gp130 did not appear to be significantly altered between delayed and active blastocysts (Hamatani et al., 2004).

1.3 Germ cell development

1.3.1 Germ cell specification

In mice, and possibly all mammals, the initiation of the germline depends on secreted signals from the previously segregated extraembryonic lineages (Hayashi et al., 2007). Inducing signals from the Exe and VE, like *Bmp4*, *Bmp8b* and *Bmp2*, instruct a small number of adjacent pluripotent proximal epiblast cells to become PGCs while the

majority of epiblast cells differentiate towards a somatic cell fate (Saitou, 2009). It appears that until this stage epiblast cells are not committed to develop into a certain cell fate as transplantation of epiblast cells from the distal tip of the embryo next to the Exe showed that those cells were capable to develop into PGCs when exposed to the inducing signals (Tam and Zhou, 1996). Mutations in either of those signaling genes cause failure of PGC specification resulting in either absence or severe reduction of PGCs (Lawson et al., 1999; Ying et al., 2000; Ying and Zhao, 2001). In order to analyse the difference between epiblast cells that will commit to a PGC and those that will commit to a somatic cell fate Saitou et al. (2002) investigated the transcriptional profile of epiblast cells at single cell resolution. They found that *Bmp4* signaling induced expression of *Fragilis*, an interferon-induced transmembrane protein, in epiblast cells, that acquire germ cell competence (Saitou et al., 2002). The formation of lineage restricted PGC precursors is marked by expression of *Blimp1* in six of the *Fragilis* positive cells which are in direct contact with the overlying Exe (Ohinata et al., 2005; Hayashi et al., 2007). The function of *Blimp1* during germ cell specification involves the suppression of the somatic cell fate, expressed by their non-PGC neighbours, and the induction of germ cell character in the *Fragilis* positive epiblast cells. Germ cell character is induced by repression of mesodermal genes like *Hoxa1 and Hoxb1*, both of which are expressed to a high level in the surrounding somatic cells and upregulation of pluripotency-associated genes such as *Nanog* and *Stella* (Ohinata et al., 2005). *Blimp1*$^{-/-}$ embryos initially form a cluster of about 20 PGCs but fail to repress somatic cell lineage specific gene expression and fail to migrate towards the developing genital ridges. By 7.25dpc a cluster of 40 *Stella* (earliest known marker for fate-restricted PGCs) and alkaline phosphatase (marker of an undifferentiated state) positive founder PGCs can be observed at the posterior end of the primitive streak in the extraembryonic mesoderm (Exm) (Saitou, 2009). Around 7.5dpc PGCs initiate migration where they travel from the extraembryonic tissues back into the embryo, passing through the developing hindgut, before reaching and colonising the emerging genital ridges at 10.5 dpc (Bendel–Stenzel et al., 1998). During this time, the number of PGCs increases to more than 3,000 and by 13.5 dpc, the number in the fetal gonads reaches around 22,000 (Tam and Snow, 1981).

1.3.2 Epigenetic remodeling during early PGC differentiation

During the migratory period PGCs undergo substantial global epigenetic changes (Surani, 2001). This is followed by widespread epigenetic reprogramming at 11.5dpc which might be expected for cells that must ensure that the zygote acquires totipotency in order to give rise to the next generation (Sasaki and Matsui, 2008). Initially, the chromatin of *Blimp1* expressing PGCs comprises similar genome-wide repressive epigentic marks, such as DNA methylation, Histone 3 lysine 9 dimethylation ($H3K9me^2$) and H3K27 trimethylation ($H3K27me^3$), like the chromatin of their surrounding somatic neighbors (Seki et a., 2007; Sasaki and Matsui, 2008). Migrating PGCs initiate global changes in DNA methylation, $H3K9me^2$ and $H3K27me^3$ (Seki et al., 2005). DNA methylation and $H3K9me^2$, which are generally associated with long-term rather than transient repression, are gradually erased resulting in a global chromatin state that is largely free of repressive chromatin marks between 7.5 and 8.25dpc (Seki et al., 2005). This coincides with transient inhibition of RNA polymerase II dependent transcription, possibly to circumvent transcriptional misregulation caused by the reduction in repressive epigenetic marks (Seki et al., 2007). From approximately 8.25dpc there is an increase in global $H3K27me^3$ levels, a less permanent repressive mark, that also correlates with the erasure of $H3K9me^2$. This possibly occurs in order to inhibit the expression of lineage specific genes that are upregulated during early differentiation and to maintain an appropriate repressive chromatin state of the PGC genome. Since this mark is considered a less permanent repressive mark than $H3K9me^2$ it is possible that its role also involves keeping lineage specific silenced genes poised for rapid transcriptional activation and helps to maintain pluripotency. Furthermore, in female germ cells reactivation of the X-chromosome is also initiated during the migratory period of PGCs (de Napoles et al., 2007; Sugimoto et al., 2007). By 10.5dpc, the PGCs have migrated to the genital ridges and undergo further large-scale chromatin remodeling that results in genomic imprints being erased (Chong, and Whitelaw, 2004). Genomic imprinting itself refers to a genetic phenomenon by which certain genes are expressed in a parent-of-origin-specific manner. Deletion of imprints is reflected by

demethylation of the imprinted loci (Chong, and Whitelaw, 2004). This occurs simultaneously with demethylation of other regions of the PGC genome (Sasaki and Matsui, 2008). Once parental imprints have been reset new, gender specific imprints must be introduced (Sasaki and Matsui, 2008). Imprints are re-established after sex determination has been initiated and germ cells have differentiated to become either male or female in order to give rise to sperm or oocytes, respectively (Sasaki and Matsui, 2008).

1.3.3 Sex determination

Before colonising the prospective gonadal tissues XX and XY PGCs appear to behave identically in most aspects. The germ cells' decision to embark on either a male or female developmental pathway appears to be determined through the embryonic gonadal environment, rather than the chromosomal sex of the germ cells themselves (Kocer et al., 2009). Indeed, it has been shown that XY germ cells in female chimaeric embryos can develop as oocytes and XX germ cells are able to differentiate into prospermatogonia in male chimaeric embryos (Ford et al., 1975; Palmer and Burgoyne, 1991). In contrast, the decision of the initially bipotential, undifferentiated gonad to develop along a female (ovarian) or male (testicular) pathway depends on the inheritance of X and Y sex chromosomes (Matzuk and Lamb, 2008; Kocer et al., 2009). Differentiation of the primordial gonad into either ovary or testis ultimately directs the sexual development of the rest of the embryo (Kocer et al., 2009). The decision whether gonadal somatic cells differentiate into testicular sertoli cells or ovarian granulosa cells is controlled by the presence or absence of the Y chromosomal testis determining factor *Sry* between 10.5dpc and 12.5dpc (Lovell-Badge and Robertson, 1990; Koopman et al., 1991; Sekido et al., 2004*).* Expression of *Sry* upregulates *Sox9* which appears to be both necessary and sufficient to induce testicular development. XY females are often characterised by mutations that inactivate *Sry*. Ectopic expression of this gene causes testis formation and physical and behavioral sex reversal in XX embryos (Gubbay et al.,

1990; Sinclair et al., 1990; Koopman et al., 1991; Cameron and Sinclair, 1997). Similarly expression of a *Sox9* transgene in a XX mouse promotes testis development of XX gonads (Bishop et al., 2000; Vidal et al., 2001). *Wnt/b*-catenin signaling in the XX gonadal somatic cells ensures, at least in part, female sex determination by antagonising *Sry/Sox9* signaling (Kim et al., 2006; Chassot et al., 2008). This is supported by the observation that ectopic expression of a stabilised form of beta-catenin in 11.5 dpc gonadal somatic cells results in ovary formation in XY mice (Maatouk et al., 2008).

In the adult gonads expression of the transcription factors *Foxl2* in granulosa cells and *Dmrt1* in sertoli cells appears to be essential to maintain sex determination in mice long after the choice to embark along a female or male developmental pathway in the embryo has been made. A mutation in *Foxl2* in the adult mouse can cause transdifferentiation of granulosa cells towards a sertoli cell fate and loss of *Dmrt1* appears to upregulate *Foxl2* and causes transdifferentiation of sertoli cells into granulosa cells in adult testis (Uhlenhaut et al., 2009; Matson et al., 2011).

As mentioned above a germ cells' decision to develop as either male or female is guided by sex-determining signaling molecules secreted by the prospective testes or ovaries, respectively (Kocer et al., 2009). Initially, between 12.5 to 13.5dpc male and female embryonic gonads are colonised by post-mitotic/pre-meiotic germ cells which represent the transition between the end of mitosis and the start of meiosis (Hilscher et al., 1974; McLaren, 1984; McLaren, 2003). Work from Adams and McLaren (2002) suggests that germ cells in XY gonads have responded to the testicular environment and are committed to spermatogenesis by 12.5dpc. Germ cells in XX gonads appear to have responded to the ovarian environment by 13.5dpc and are fully committed along a female pathway (Adams and McLaren, 2002). The first sex-specific morphological difference between male and female germ cells becomes evident when female germ cells enter meiosis at 13.5dpc.

1.3.4 Meiosis in males and females

Between males and females the general outline of meiosis (described under 1.4.2) is conserved but the details are remarkably different. One sex specific difference is in the timing of meiosis (figure 1.6). In both sexes PGCs continue to divide mitotically until about 13.5 dpc (McLaren and Southee, 1997). In a male embryo the germ cells arrest in the G0 stage of the cell cycle, resuming mitosis after birth (Hilscher et al., 1974; McLaren, 1984). Meiosis begins at 8 to 10 dpc and continues without interruption throughout the life of the male due to a continuous supply of diploid spermatogonia entering meiosis. In contrast, female meiosis begins during fetal development and all oocytes that the female will possess in her lifetime are produced at this stage. For this process germ cells stop proliferating at 13.5dpc and enter prophase of meiosis I as oocytes. The oocytes pass through the different stages of prophase and then arrest at diplotene around the time of birth. This specialized prolonged arrest state is known as dictyate (Speed, 1982). During the oocytes' meiotic arrest, homologous chromosomes are physically held together as bivalents by crossover events and cohesion complexes between the deoxyribonucleic acid (DNA) molecules. Oocytes remain arrested in this state until sexual maturation; meiosis is not continued until hormonal stimulation shortly before ovulation and the second division is completed when the egg is fertilized (figure 1.4) (McLaren, 2003).

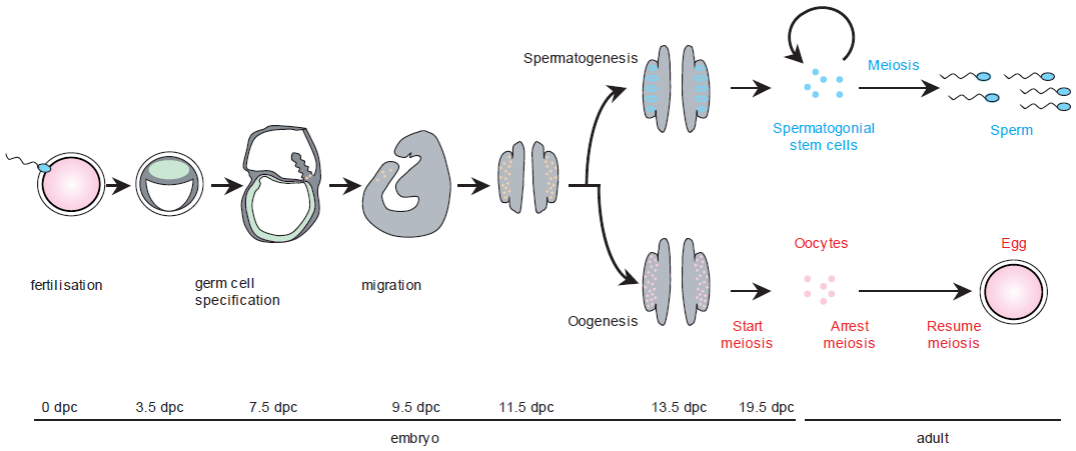


Figure 1.5 Meiosis in males and females. Pluripotent cells (green) are found in early mouse embryos. By 7.5dpc primordial germ cells (PGCs, orange) have differentiated from the pluripotent cells. The PGCs proliferate and migrate towards the genital ridge. They have colonized the developing gonads by 11.5 dpc. Around 13.5 dpc PGCs in male embryos enter quiescence. A few days after birth some of the male germ cells differentiate into spermatogonial stem cells. Those provide a continuous supply of spermatogonia that undergo meiosis and differentiate into mature sperm in the adult mouse. In contrast PGCs in female embryos initiate meiosis at around 13.5 dpc but arrest at the end of prophase I a few days after birth. Meiosis resumes in fully grown adults upon ovulation and the second meiotic division is only completed upon fertilization, taken from Öllinger et al. (2010).

1.4 Mitosis and Meiosis

Meiosis is a specialized type of cell division during which the number of chromosomes is reduced by half so that fertilization during sexual reproduction results in the generation of a diploid karyotype. The reduction of chromosome number is accomplished by essentially one round of DNA replication followed by two rounds of cell division, meiosis I and meiosis II. This process creates haploid daughters from a diploid parental cell in a manner that ensures each daughter cell a complete haploid genome (Page and Hawley, 2003). In order to understand how chromosomes are segregated during meiosis it is important to be familiar with the principles of chromosome segregation during mitosis.

1.4.1 Mitosis

Mitosis is a form of eukaryotic cell division which results in the production of two daughter cells genetically identical to the parent cells and each another. Mitosis is preceeded by interphase and then subdivided into prophase, prometaphase, metaphase, anaphase and telophase during which the genetic information of a cell is first duplicated and then segregated into the two daughter cells (outlined in figure 1.6) (Sullivan and Morgan, 2007). During interphase the cell prepares for cell division by increasing in size and replicating its DNA. DNA replication results in two identical copies of each chromosome which are then called sister chromatids. During mitosis sister chromatids need to be accurately segregated to ensure accurate transmission of genetic material to each daughter cell. A key difference between mitosis and meiosis is that sister chromatids remain joined after metaphase in meiosis I, whereas in mitosis they separate (figure 1.7).

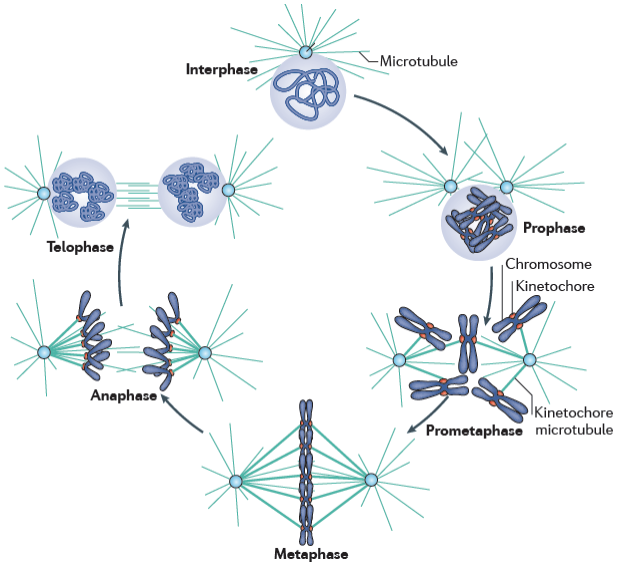


Figure 1.6 Mitotic chromosome segregation. The various stages of the cell cycle are depicted. During interphase, the cell undergoes growth and replication of the DNA. Upon replication of the spindle pole body and DNA, the cell undergoes a second round of growth and subsequently enters mitosis. Mitosis is divided into prophase (when the chromatin is condensed), prometaphase (when kinetochore microtubules start to interact with kinetochores), metaphase (when chromosomes become bi-oriented), anaphase (when the sister chromatids segregate to opposite spindle poles) and telophase (when chromosomes decondense). In most eukaryotes, the nuclear membrane degrades during mitosis and reforms during telophase, but this does not occur in budding yeast. Figure was taken from Verdaasdonk and Bloom, (2011).

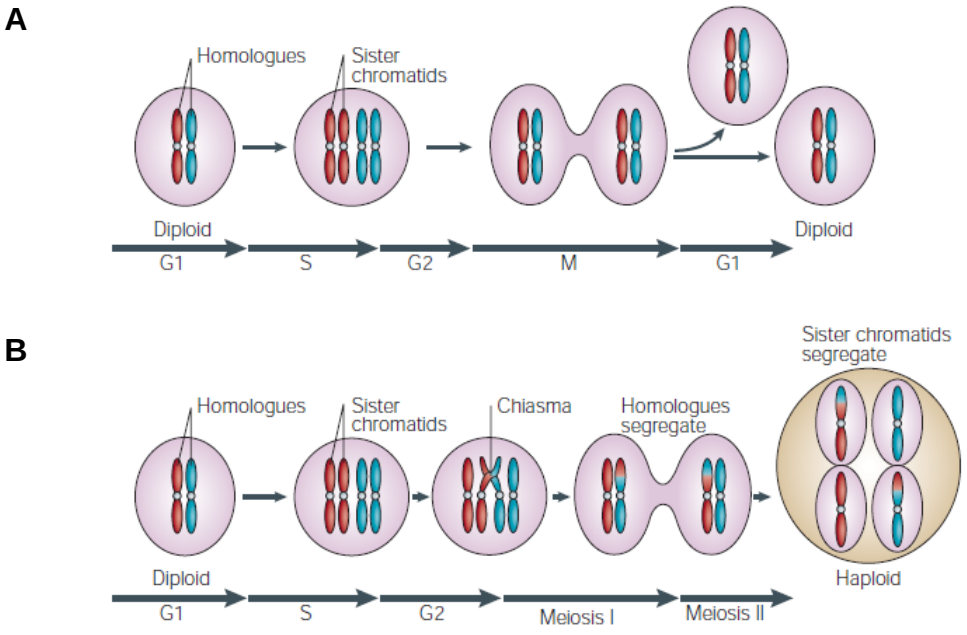


Figure 1.7 Mitosis and Meiosis (A) Interphase which preceeds mitosis is subdivided into G1, S and G2 phase. During G1 synthesis of various enzymes that are required in S phase is taking place. During S phase diploid cells replicate chromosomes. This is followed by the G2 phase where microtubules and other proteins are produced that are required during the process of mitosis. During mitosis (M) sister chromatids segregate so that diploid daughters are produced. (B) During meiosis two rounds of chromosome-segregation (meiosis I and II) follow a single round of DNA replication during pre-meiotic S-phase. Meiosis I results in the segregation of homologous chromosomes (shown in red and blue), which are held together by chiasmata, are segregated to opposite poles. During meiosis II sister chromatids get segregated resulting in the formation of haploid gametes taken from Marston and Amon (2004).

In order for sister chromatids to be correctly segregated bi-orientation on the mitotic spindle needs to be achieved. Bi-orientation refers to a process where sister chromatids align on the mitotic spindle in a manner that allows microtubules emanating from opposite spindle poles to attach to the sister chromatid kinetochores in order to pull them apart to opposite spindle poles (see figure 1.5) (Verdaasdonk and Bloom, 2011). Attachment of microtubules to sister chromatid kinetochores can be erroneous probably because of the 'search and capture' nature of this process (Musacchio and Salmon, 2007). In order to ensure fidelity of chromosome segregation formation of bi-orientation is monitored by the spindle-assembly checkpoint (SAC). The SAC delays anaphase if

chromosomes are not properly attached so that the error can be corrected and therewith ensures faithful chromosome segregation in mitosis (Musacchio and Salmon, 2007). The linkage between sister chromatids brings the two centromeres/kinetochores into close proximity which is important for correct microtubule attachment (Shintomi and Hirano, 2010). Sister chromatids are held together from DNA replication until the onset of chromosome segregation by the cohesin complex. The cohesin complex consists of four core subunits: two subunits of the structural maintenance of chromosomes (SMC) protein family, Smc1 and Smc3; the kleisin family protein Scc1/Rad21 and Scc3/Psc3 (Nasmyth and Haering, 2005). The cohesin proteins associate with the chromosomes before DNA replication and are afterwards converted into the physical linkage that holds the sister chromatids together through G2 phase. It has been shown that in vertebrates the major bulk of cohesin disassociates from the sister chromatid arms during prophase, before the sisters are separated (Losada et al. 1998; Waizenegger et al. 2000). Dissociation of the cohesin complex from the chromatid arms might involve phosphorylation of cohesin subunits as several residues of the cohesin subunits SA2/Scc3 and RAD21/Scc1 are phosphorylated in a mitosis-specific manner. Furthermore, expression of an only poorly phosphorylatable mutated version of SA2/Scc3 in HeLa cells results in inefficient dissociation of cohesin from chromatin during prophase I (Hauf et al., 2005). This suggests that phosphorylation of cohesin subunits enables the cohesin complex to dissociate from the chromatin during prophase. Despite the dissociation of cohesin from the chromatid arms during prophase I sister chromatid cohesin is maintained until the onset of anaphase. This is facilitated by centromeric cohesin which escapes release from chromatin during prophase (Waizenegger et al. 2000). It has been suggested that centromeric cohesin is protected from phosphorylation by Shugoshin (*Sgo*) to ensure that cohesin persists at the centromeres until anaphase (Shintomi and Hirano, 2009). This theory is based on the observation that the Sgo protein is found to localise to the centromeres during prophase and disappears from the centromeres at the onset of anaphase. Knockdown of *Sgo* by RNA interference (RNAi) in HeLa cells causes premature separation of sister chromatids resulting in cell cycle arrest (McGuinness et al., 2005). At metaphase the

nuclear envelope has disappeared (the fragmentation of which starts during prophase, and is complete in pro-metaphase), and the sister chromatids align on the cells equator. This is followed by the onset of anaphase, during which a specific endopeptidase, called separase, is activated to cleave the kleisin subunit Scc1/Rad21 (Uhlmann et al., 1999). This results in disassociation of the cohesin complex, thereby allowing chromatid separation. The chromatids then move to opposite ends of the cell towards the spindle poles (Sullivan and Morgan, 2007). Anaphase is followed by telophase which results in reassembly of the nuclear envelope around each daughter nuclei. At the same time the binucelate cell is divided into two by the process of cytokinesis to form the two daughter cells with each daughter possessing a full copy of the genome of its parent cell mitosis is complete.

1.4.2 Meiosis

Chromosome segregation during meiosis depends to some extent on the same machinery used during mitosis. However, since meiotic cells reduce chromosome numbers by undergoing two rounds of chromosome segregation after only one round of DNA replication, several meiosis-specific innovations are required to generate haploid gametes (Handel and Schimenti, 2010; Öllinger et al., 2010). During the first round of chromosome segregation, meiosis I, only one of the two homologues of each duplicated chromosome is partitioned into each of the two daughter cells. During the second round of chromosome segregation, meiosis II, the duplicated chromosomes separate into their constituent sister chromatids which are then partitioned equally between two daughter cells. Prior to entering meiosis, the DNA is replicated during premeiotic S-phase that converts each chromosome into two sister chromatids resulting in a cell with four chromatids of each type of chromosome – two maternal and two paternal chromatids (Handel and Schimenti, 2010; Öllinger et al., 2010). Both the two maternal and paternal sister chromatids are referred to as homologue chromosomes. Premeiotic DNA replication is followed by prophase I, which is subdivided into leptotene, zygotene, pachytene, diplotene and diakenisis, during which homologues chromosomes pair, align

and undergo recombination which will result in crossover (CR) formation between the DNA molecules of homologous chromosomes (Jones, 2008). The CR events physically connect homologue chromosomes until metaphase I. The first meiotic division segregates homologous chromosomes rather than sister chromatids and therefore a physical connection between homologous chromosomes is required to allow proper bi-orientation of homologous chromosomes on the meiotic spindle and subsequent segregation to opposite spindle poles (Schvarzstein et al., 2010). Connection between sister chromatids is maintained until metaphase II when the second meiotic division segregates sister chromatids to opposite spindle poles in order to generate haploid gametes.

1.4.2.1 Sister chromatid cohesion

During meiosis sister chromatids are paired and physically bound together by incorporation of meiosis specific cohesion proteins (Uhlmann and Nasmyth, 1998). This is initiated during premeiotic DNA replication when the mitotic cohesin subunit Scc1/Rad21 is largely replaced by the meiosis specific subunit Rec8 (Eijpe et al., 2003). During the leptotene stage of prophase I the remaining meiotic cohesin components, Smc1β, Smc3, and Stag3, associate with the meiotic chromosomes. Work in fission yeast and mammals showed that mutations in meiotic cohesin proteins not only results in premature sister chromatid separation but also defects in meiotic recombination and synaptonemal complex formation (Klein et al., 1999; Eijpe et al., 2000; Pelttari et al., 2001; Eijpe et al., 2003; Revenkova et al., 2004; Xu et al., 2005). This data, together with further work carried out in fission yeast which showed that ectopic expression of Scc1/Rad21 is not able to restore meiotic recombination and synaptonemal complex formation in *rec8*Δ cells suggests that meiotic-specific cohesin is not only required to hold sister chromatids together, but also to generate the correct chromosomal architecture for the subsequent meiotic processes that are discussed below (Toth et al., 2000; Yokobayashi et al., 2003).

1.4.2.2 Meiotic recombination

Meiotic recombination is initiated by the formation of DNA double-strand breaks (DSBs) which are introduced into the genome by the meiosis-specific endonuclease Spo11 (Romanienko and Camerini-Otero, 2000). The DSB are then resected on either side of the break by 5'–3' exonucleolytic processing to create 3' overhangs of single-strand DNA (ssDNA) which are initially bound by replication protein A (RPA) (Symington and Gautier, 2011). In mammals the DSB resection is dependent on a variety of factors including the MRN complex, CtIP, Sgs1 and Exo1 (Bernstein and Rothstein, 2009; Youds and Boulton, 2011). Replacement of RPA by Rad51 recombinase and its meiosis-specific paralogue Dmc1 marks the assembly of early recombination foci (Li and Ma, 2006). Processing of the DNA double-strand breaks at these early recombination foci is crucial for homologue chromosomes to pair. *Spo11*$^{-/-}$ and *Dmc1*$^{-/-}$ mice are infertile as they fail to pair and synapse homologues chromosomes which possibly reflects the involvement of early meiotic recombination in the homology search (Pittman et al., 1998; Baudat et al., 2000). Association of DSBs with Rad51 and Dmc1 results in the formation of nucleoprotein filaments which invade adjacent DNA molecules to search for homologous DNA sequences (Hunter and Kleckner, 2001). In order to generate exchanges between homologous chromosomes maternal filaments have to invade paternal chromatids and vise versa as invasion of sister chromatids would, like it takes place in mitotic cells, simply result in repairing the DSB (Collins and Newlon, 1994; Schwacha and Kleckner, 1994; Paques and Haber, 1999). The mechanisms that biases invasion towards homologous chromosomes rather than sister chromatids during meiosis are currently not well understood but likely require several meiosis specific proteins (Schwacha and Kleckner, 1997). For example, *Rec8* null mice *(Rec8*$^{-/-}$*)* show recombination between sister chromatids instead of homologous chromosomes (Xu et al., 2005). This suggests that *Rec8* plays a role in directly limiting synapsis to homologous chromosomes during mammalian meiosis or organising a meiotic chromosome structure which favours recombination between homologous chromosomes. Recombination between homologous chromosomes proceeds when the

invading 3' filament of one of the homologues becomes paired with a complementary strand of the other homologous chromosome to create recombination intermediates (Petronczki et al., 2003). Interaction between these invasion intermediates and the second DNA double strand break end results in the formation of a meiotic recombination intermediate known as a double Holliday junction that possesses regions of heteroduplex DNA containing complementary DNA strands from the two homologous chromosomes (Youds and Boulton, 2011). The double Holliday junctions are then subjected to further processing to be resolved as either as interhomolog crossovers (CRs) or as non-crossovers (NCRs) (see figure 1.8) (Constantinou et al., 2001; Öllinger et al., 2010; Youds and Boulton, 2011). A NCR event refers to a situation where only small regions of single DNA strands have been exchanged between chromatids from homologous chromosomes whereas in the case of a CR event both DNA strands of the recombining chromatids have been exchanged between homologous chromosomes at the recombination site. CR and NCR events are marked by loss of early recombination markers, like Dmc1 and Rad51, from meiotic chromosomes and as the recombination sites mature they incorporate the recombination-associated proteins Msh4 and Msh5 (Youds and Boulton, 2011). Data from *Msh4*$^{-/-}$ and *Msh5*$^{-/-}$ mice suggests that both proteins are crucial for the formation of CRs. Mice with a mutation in either protein fail to form CRs and are consequently sterile (Edelmann et al., 1999; Kneitz et al., 2000). As meiosis proceeds, the majority of recombination sites in the nucleus is resolved as NCRs as shown by a decrease of >200 early recombination events to around twenty-five Mlh1-containing CRs (Moens et al., 2007). Mlh1, which participates in the late step of CR formation, is considered a marker for recombination that will mature as CRs. This is based on the fact that Mlh1 appears to exclusively associate with recombination sites that are designated to form CRs (Broman et al., 2002; Froenicke et al., 2002). Mature CR events appear as chiasmata and as they are involved in maintaining the physical connection between homologous, which ensures two homologues are being pulled to opposite poles at meiosis I. It is essential that every chromosome has at least one CR in order to achieve faithful chromosome segregation to take place during the first meiotic division (Carpenter, 1994). This is well demonstrated by the phenotype of *Mlh1*$^{-/-}$ mice

which show defective chiasmata formation resulting in meiotic arrest (Edelmann et al., 1996). The mechanisms that regulate the number and distribution of chiasmata in the meiotic genome are still not clear. However, recent evidence suggests that the location of DSB formation might be influenced by chromatin structure and certain sequence motifs (Fukuda et al., 2008; Mets et al., 2009; Baudat et al., 2010). Figure 1.8 summarises the processes of CR and NCR formation.

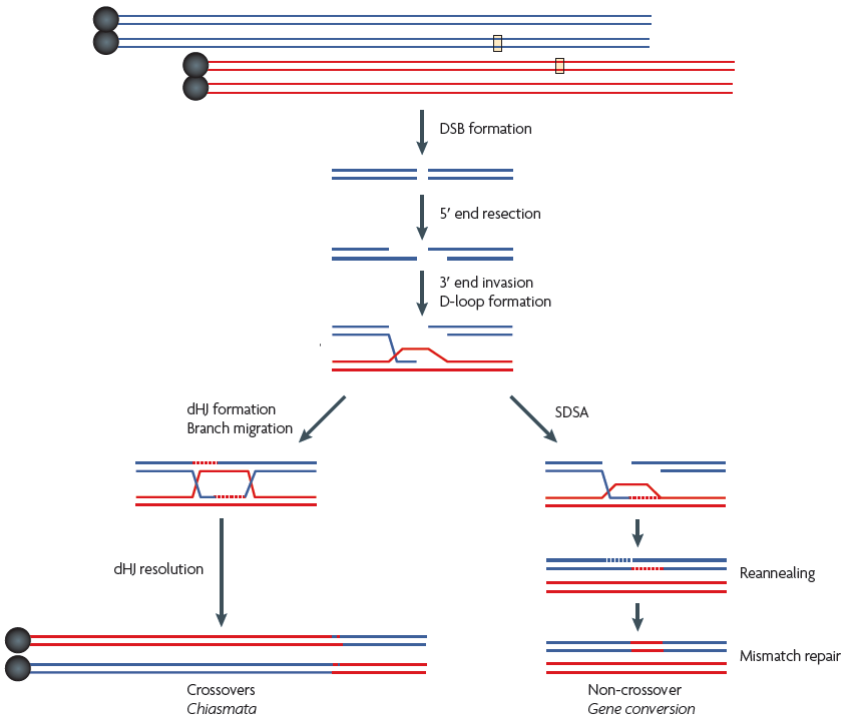


Figure 1.8 Model for meiotic crossover or non-crossover formation. Red and blue lines represent single DNA strands. Generated DSBs are resected to generate ssDNA with 3′ overhang. Strand invasion events result in recombination intermediates. If the second end of the original DSB also engages with the homologue, a double Holliday junction is formed which leads to crossover products (left hand side). The majority of recombination intermediates are not resolved as crossovers but as non-crossovers. Non-crossovers are formed when the invading strand associates with the opposite end of the original break, as in synthesis-dependent strand annealing (SDSA). Figure was taken from Handel and Schimenti (2010).

1.4.2.3 Meiotic chromosome synapsis

In order for homologous recombination between chromosomes to initially occur, homologous sequences must first find each other. How exactly each chromosome searches for and identifies its homologous partner for recombination is not well understood but it is thought that the interaction of the meiotically induced DNA DSBs (described previously) with matching sequences on the homologous chromosome, brings the axis of homologous chromosomes into alignment (Page and Hawley, 2004). Homologue chromosomes roughly align their axes with their partner during the zygotene stage of meiosis and chronologically with homologous chromosome alignment are chromosome pairing and synapsis taking place. During homologous chromosome synapsis the synaptonemal complex (SC) assembles between homologous chromosomes which is thought to function as a scaffold to enable the repair and recombination processes and ultimately crossover formation between interacting chromatids as described above (Page and Hawley, 2004; Costa and Cooke, 2007). The SC itself is a meiosis-specific tripartite proteinaceous synaptonemal complex structure comprised of axial elements (AEs), transverse filaments (TFs) and the central element (CE) (Page and Hawley, 2004; Costa and Cooke, 2007). Assembly of the synaptonemal complex is initiated during leptotene when the axial element (AE) proteins Sycp2 and Sycp3 assemble on the meiotic chromosomal axes (Schalk et al. 1998; Pelttari et al. 2001). This is followed by the assembly of TFs, mainly consisting of Sycp1, and CE proteins Syce1, Syce2 and Tex12 during zygotene (Costa et al., 2005; Öllinger et al., 2005; Hamer et al., 2006). Formation of TFs and the CE facilitates close connections between the AE associated with one pair of homologous sisters and with the AE associated with the other pair. The AEs from this stage onwards are termed lateral elements (LEs) (Costa and Cooke, 2007). The completely assembled SC, which has formed at pachytene, physically links homologous chromosomes in a zipper-like manner and thus mediates synapsis (Öllinger et al., 2010).

Studies from *Sycp2*$^{-/-}$ and *Syp3*$^{-/-}$ mice have shown that assembly of the AEs during leptotene is required for subsequent formation of the SC during spermatogenesis. Male *Sycp2*$^{-/-}$ and *Sycp3*$^{-/-}$ mice show impaired meiotic chromosome synapsis as Sycp1 filaments, which are generally considered to mark synapsis, neither assemble nor function properly in the absence of either protein. Immunofluorescence and electron microscopy showed that only short and fragmented stretches of Sycp1 are associated with the meiotic chromosomes (Yuan et al., 2000; Yang et al., 2006). In contrast *Sycp2*$^{-/-}$ and *Sycp3*$^{-/-}$ females show, based on Sycp1 localisation, more complete synapsis and are fertile albeit at a reduced level (Yuan et al., 2002; Yang et al., 2006). This reduction in fertility in *Sycp3*$^{-/-}$ females is caused, at least in part, by a failure in CR formation resulting in high rates of aneuploidy in mature oocytes and subsequently embryo lethality (Yuan et al., 2002). The phenotype of *Sycp1*$^{-/-}$ mice demonstrates that Sycp1 is essential for the formation of homologous chromosome synapsis. AEs assemble and align normally but synapsis does not take place in the absence of Sycp1 resulting in sterility of both sexes (de Vries et al., 2005). Furthermore, it was shown that late Mlh1 containing recombination foci are not formed in *Sycp1*$^{-/-}$ spermatocytes resulting in impaired CR formation (de Vries et al., 2005). Similarly mutations in AE and TF components disrupt homologous chromosome synapsis. In the absence of either of the CE proteins short regions of synaptonemal complex assemble but are unable to extend along the entire chromosome axis and early recombination events do not mature into meiotic CRs resulting in male and female sterility (Bolcun-Filas et al., 2007, Hamer et al., 2008; Bolcun-Filas et al., 2009). In addition to components of the SC, meiotic specific cohesin proteins facilitate the completion of synapsis between homologous chromosomes (Revenkova et al. 2001; Eijpe et al. 2003; Revenkova et al. 2004). For example *Smc1β-/-* mice exhibit high levels of incomplete synapsis which is thought to be caused by defects in the compaction of meiotic chromatin (Revenkova et al. 2004). Taken together it appears that the events of cohesin protein assembly, meiotic recombination, and homologous chromosome pairing and synapsis during meiosis are closely interdependent. Homologous pairing and chromosome synapsis appear to be initiated by meiotic recombination which in turn allows meiotic recombination sites to

mature. Pachytene is followed by the diplotene stage of meiosis during which the synaptonemal complex disassociates (Costa and Cooke, 2007). Connection between homologous chromosomes at this stage is maintained by the previously mentioned chiasmata and by cohesin complexes linking the sister chromatid arms (Revenkova and Jessberger, 2005). These homologous chromosome pairs, called bivalents, can then align on the meiotic spindle for reductional segregation in meiosis I.

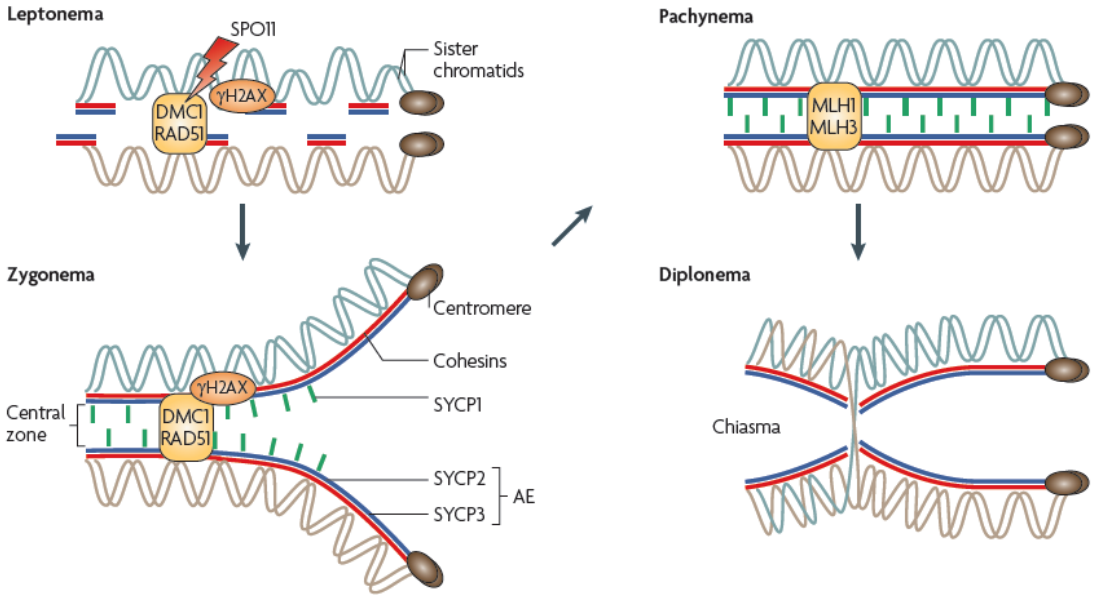


Figure 1.9 Overview of the processes occurring during prophase I of meiosis. During the leptotene (leptonema) stage of prophase I homologous chromosomes begin to align and AEs begin to assemble on sister chromatids of each homologous chromosome (turquoise and gold lines). Spo11 induces DSBs, characterised by γH2AX accumulation, which mediates recombination. Formation of early recombination foci is marked by binding of Dmc1 and Rad51 to DSBs. During zygonema (zygotene) chromosomes have been aligned and synapsis is initiated as demonstrated by formation of the TFs (here shown by assembly of Sycp1). Formation of the SC is complete by pachytene (pachynema). During pachytene a small proportion of recombination intermediates matures into crossovers as marked by assembly of Mlh1 and Mlh3. When recombination is completed the SC disassembles and homologous are held together by chiasmata (this stage is referred to as diplonema or diplotene). Figure was taken from Handel and Schimenti (2010).

1.4.2.4 Segregation of homologous chromosomes and sister chromatids during meiosis

Metaphase I is preceded by diakinesis, the last stage of prophase I, which is very similar to pro-metaphase of mitosis, as during this time the nuclear membrane disintegrates and the meiotic spindle begins to form (Schvarzstein et al., 2010). The physical connection between homologous chromosomes achieved by chiasmata and cohesin proteins is essential to achieve bi-orientation of homologous chromosomes on the meiotic spindle midway between the poles and subsequent traction of maternal and paternal kinetochores toward opposite poles during metaphase I (Hauf and Watanabe, 2004). To ensure that homologous chromosomes, rather than sister chromatids, segregate during meiosis I the sister chromatids of each homologous chromosome have to behave as a single functional unit and attach to microtubules emanating from the same pole (Sakuno and Watanabe, 2009). It has been suggested that during MI, similar to mitosis, a SAC functions to delay anaphase until proper kinetochore-microtubule attachments are formed (Wassmann et al., 2003; Niault et al., 2007; Hached et al., 2011). Once correct bi-orientation is achieved cohesin proteins are lost from chromosome arms during the transition from metaphase I to anaphase I but not the centromeres as shown by immunostaining for Rec8, Smc3 and Smc1β on meiotic chromosomes (Revenkova et al. 2001; Eijpe et al. 2003). Displacement of cohesin complexes from the chromosome arms requires separase-mediated proteolytic cleavage of the cohesin subunit Rec8 (Kudo et al., 2006). Dissociation of the cohesin complex from the chromosome arms is thought to allow the chiasmata that link the bivalent to resolve and the homologous chromosomes to separate during the metaphase-anaphase transition of meiosis I. The persistence of cohesin complexes at the centromeres ensures that the physical connection between sister chromatids is maintained at this stage (Schvarzstein et al., 2010).

In mammals *Shugoshin-like-2* (*Sgol2*) is responsible to protect centromeric cohesin until the metaphase II to anaphase II transition. It has been shown that *Sgol2*$^{-/-}$ mice are unable to retain centromeric cohesin during meiosis I resulting in single chromatids at

metaphase II which will give rise to aneuploid gametes resulting in infertility (Lee et al., 2008; Llano et al., 2008). Meiosis II appears similar to mitotic chromosome segregation as it results in separation of sister chromatids. During meiosis II the connection between sister chromatids, mediated by centromeric cohesin, facilitates bi-orientation and the attachment of sister kinetochores to microtubules emanating from opposite poles of the spindle (Hauf and Watanabe, 2004). During the metaphase-anaphase transition of meiosis II *Sgol2* becomes inactivated and sister chromatid separation is mediated by cleavage of the remaining Rec8 by separase (Kudo et al., 2006). The sister chromatids by convention are now called sister chromosomes. The last step of meiosis is telophase II, which is similar to telophase I, and is marked by de-condensing and lengthening of the chromosomes as well as the disappearance of the spindle and reassembly of the nuclear envelope around each daughter nuclei (Strachan and Read, 2004). Meiosis concludes with the production of four haploid spermatozoa in males and two haploid ova in females which completes meiosis. The meiotic processes described in this, and the previous section are summarised in figure 1.10.

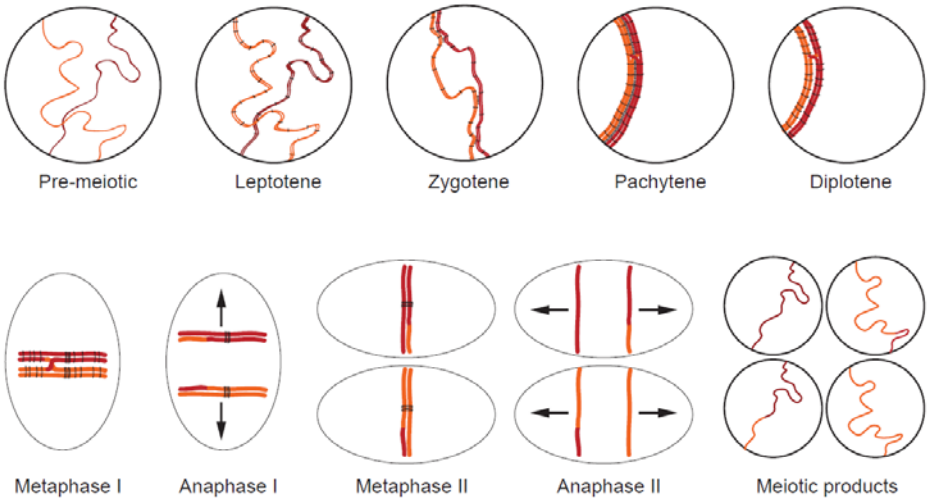


Figure 1.10 Chromosomal behaviour during mouse meiosis Pre-meiotic germ cells contain one maternal and one paternal copy of each chromosome which are also called homologous copies or chromosomes (orange and red threads, only two homologous chromosome axes are shown for clarity). Prior to entering meiosis DNA replication duplicates each chromosome. The duplicated chromosomes are referred to as sister chromatids and are held together by meiosis-specific cohesion complexes (black rings). During leptotene the chromosomes are condensing and meiotic recombination is initiated (not shown) and the axial elements of the synaptonemal complex assemble along the chromosomal axes (not shown). This is followed by zygotene where homologous chromosomes start to pair and then synapse as the synaptonemal complex continues to assemble (black zipper-like structure). Synapsis is complete by pachytene and recombination between homologous chromosomes has generated crossover sites. During diplotene, the synaptonemal complex disassembles and the crossover sites mature into chiasmata which act as a physical connection between homologous chromosomes. Homologous chromosomes align on the meiotic spindle during metaphase I. This is followed by cleavage of cohesin subunits along the chromosome arms which allows the chiasmata to resolve and homologous chromosomes to segregate during anaphase I. Cleavage of centromeric cohesions in meiosis II then facilitated sister chromatid segregation. Each round of meiosis generates four genetically distinct haploid products from a diploid cell, taken from Öllinger et al. (2010).

1.5 Spermatogenesis and oogenesis

Taken together, germ cell development ultimately aims to create developmentally competent haploid oocytes in the female and haploid sperm in the male. Fusion of oocyte and sperm results, in a diploid zygote which will undergo all of the above described developmental steps and give rise to the next generation. In order to ensure that fertilisation can occur germ cells have to, in addition to meiosis, undergo cellular differentiation and develop into mature gametes. The entire process of producing a

functional haploid oocyte or sperm is referred to as oogenesis and spermatogenesis, respectively.

1.5.1 Oogenesis

Female germ cells grow and undergo meiotic maturation within ovarian follicles (Edson et al., 2009). As described under section 1.3 germ cells in the embryonic ovary initially divide mitotically and at this stage they are referred to as oogonia. Following this, the oogonia then differentiate into oocytes and enter meiosis. Initially oogonia are arranged in clusters in the ovary. Pepling and Spradling (1998) showed by using confocal microscopy that those clusters contained 2^n germ cells as the majority of cells appear to divide synchronously. Additionally, electron microscopy revealed that the oogonia were interconnected by intercellular bridges between 11.5 to 17.5dpc caused by incomplete cytokinesis (Pepling and Spradling, 1998). This process is conserved from flies to mammals and it has been reported that defects in cluster and intercellular bridge formation causes sterility in female flies (de Cuevas et al., 1996). In mammals the function of intracellular bridges is less clear as *Tex14*$^{-/-}$ mutant female mice, which lack formation of intercellular bridges between oogonia, are fertile (Greenbaum et al., 2009). Between 20.5-22.5dpc oogonia clusters undergo repeated programme breakdown until individual oocytes remain to give rise to primordial follicles (Pepling and Spradling, 2001). A large proportion of germ cells undergo apoptosis at this stage and only about 30% of all clustered germ cells will actually form primordial follicles (figure 1.11) (Pepling and Spradling, 2001). The majority of primordial follicles are dormant until postpubertal activation and this group of primordial follicles represents the finite pool of oocytes that a female has available during her reproductive life span (Edson et al., 2009). Primordial follicles have formed by the time of birth in mice and by mid-gestation in humans, this is characterised by individual oocytes enclosed by a single layer of squamous pre-granulosa cells (figure 1.11) (Edson et al., 2009).

The function of granulosa cells involves the nurture and protection of the oocytes which in turn facilities the proliferation and differentiation of granulosa cells. Cross-talk between oocytes and somatic cells plays an important role in ovarian follicular development (McGee and Hsueh, 2000; Vanderhyden, 2002; Albertini and Barrett, 2003). For example, Dong and colleagues (1996) showed that deletion of the transcription factor *Figα* impairs formation of the pre-granulosa cell layer which results in germ cell depletion and female infertility. Similarly, mutation in the oocyte derived *growth differentiation factor-9* (*Gdf-9*) causes defects in maturation of follicular somatic cells resulting in impaired oocyte growth and maturation and hence infertility (Dong et al., 1999). In order to generate fertilisable oocytes primordial follicles are progressively activated from the resting pool from the onset of puberty. Several primordial follicles start to grow during each menstrual (in humans) or estrus cycle (in mice). Activation is irreversible and activated follicles that are not selected for further development and ovulation will undergo atresia through apoptotic cell death (McGee and Hsueh, 2000). Activated primordial follicles develop into primary follicles characterised by oocyte growth and transformation of the granulosa cells into a cuboidal shape (figure 1.11) (McGee and Hsueh, 2000). The granulosa cells proliferate and become multilayered to form secondary follicles (figure 1.11). Once the follicle is surrounded by two layers of granulosa cells an additional layer of somatic cells differentiates and forms the outermost membrane of the follicle called theca cells (Hirshfield, 1991; Tajima et al., 2007). Transition from the secondary to the tertiary or antral follicle stage involves formation of a fluid-filled cavity adjacent to the oocyte called the antrum which separates two functionally distinct granulosa cell populations (figure 1.11) (Edson et al., 2009). The granulosa cell population that surrounds the oocytes and promotes its growth, maturation and development are called cumulus granulosa cells (Matzuk and Lamb, 2002). The other granulosa cell population, mural granulosa cells, line the wall of the follicle and are involved in production and secretion of steroids as well as ovulation (Matzuk and Lamb, 2002). From the secondary follicle stage onwards granulosa cells begin to express *follicle-stimulating hormone* (*Fsh*) *receptors* and regulation of folliculogenesis switches largely from an intraovarian to extraovarian process (Edson et

al., 2009). Stimulation of granulosa cells by *Fsh* is essential for the transition from secondary follicles to antral follicles as demonstrated by infertility in *Fsh-receptor*$^{-/-}$ females which are characterised by a block in folliculogenesis before antral follicle formation (Aittomaki et al., 1995). Absence of *Fhs* causes granulosa cell apoptosis and atresia of those follicles (Aittomaki et al., 1995). Furthermore, *Fsh* stimulates granulosa cell proliferation, estrogen production, and *luteinizing hormone* (*Lh*) receptor expression (Hawkins and Matzuk, 2008). Fsh and Lh are gonadotropins secreted by the pituitary gland in response to stimulation by gonadotropin-releasing hormone (GnRH) which is produced by the hypothalamus (Hawkins and Matzuk, 2008). In response to low levels of Lh, theca cells produce androgens which are converted to estrogen in the granulosa cells (Hawkins and Matzuk, 2008). A rise in estrogen levels stimulates further development of the antral follicle and at very high levels estrogen induces an Lh surge that will finally lead to ovulation during which the oocyte is released from the ovary into the oviduct in preparation for fertilization (figure 1.11) (Edson et al., 2009). If the timing between ovulation and seminal deposition in the females' reproductive tract is correct fertilisation of the oocyte should occur successfully. Ovulation occurs in a cyclic manner during the adult life of female mammals until the pool of primordial follicles is exhausted and the female undergoes reproductive senescence.

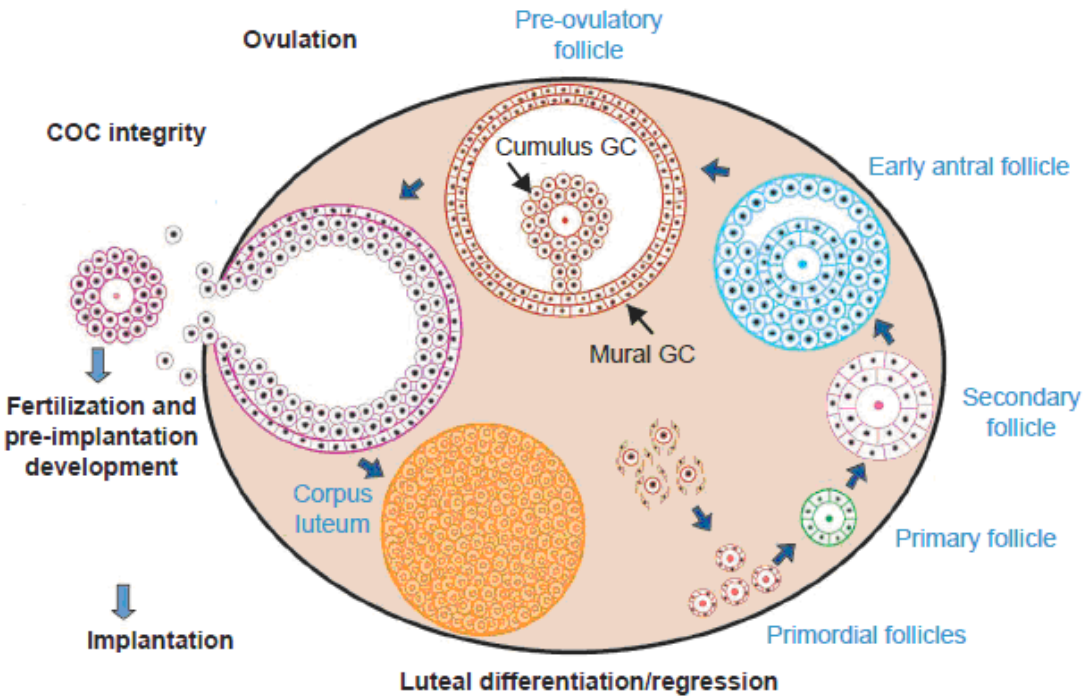


Figure 1.11 Oocyte maturation in the mammalian ovary. Clusters formed during mitotic proliferation of female germ cells will undergo programmed breakdown and individual oocytes will give rise to primordial follicles. Oocyte and granulosa cell (GC) growth and differentiation causes follicles to develop into primary, secondary and antral follicles. Ovulation releases the oocyte surrounded by cumulus cells (cumulus oocyte complex (COC)) from the ovary into the oviduct in preparation for fertilisation. Figure taken and modified from Matzuk and Lamb (2002).

1.5.2 Spermatogenesis

As described under 1.3 the decision of the initially indifferent embryonic gonad to develop into either an ovary or a testis depends on the inheritance of the X and Y chromosomes. *Sry* expression in the gonadal somatic cells induces differentiation into sertoli cells which in turn influences differentiation of other gonadal cell types along a male pathway (Palmer and Burgoyne, 1991). By 12.5dpc clusters of PGCs surrounded by Sertoli cells have organized into testis cord structures which are characteristic of XY gonads. From this point onwards PGCs are called gonocytes. The cord structures differentiate into seminiferous tubules in which the gonocytes grow and undergo meiotic

maturation to produce spermatozoa. Between 0 and 6dpp the quiescent gonocytes resume proliferation and migrate to the basement membrane of the seminiferous tubule (Huckins and Clermont, 1968; Bellve et al., 1977; de Rooij and Russell, 2000). It has been suggested that in rodents some of the gonocytes develop into differentiating spermatogonia which undergo the first round of postnatal spermatogenesis, whereas a second subpopulation of gonocytes will establishe a spermatogonial stem cell (SSC) population that then provides the basis for all subsequent rounds of spermatogenesis (de Rooij, 1998; de Rooij and Russell, 2000; Orwig et al., 2002; Yoshida et al. 2006). This theory is based on experiments which show that gonocytes, after resumption of mitotic proliferation, give rise to neurogenin 3 (Ngn3) positive and c-Kit negative cells, as well as neurogenin 3 (Ngn3) negative and c-Kit positive cells in parallel (Yoshinaga et al., 1991; Schrans-Stassen et al., 1999; Yoshida et al. 2006). Expression of Ngn3 and absence of c-Kit is characteristic for undifferentiated spermatogonia which are capable of self-renewal and the generation of differentiating cells. Upon differentiation spermatogonia downregulate Ngn3 expression and start to express c-Kit (Yoshinaga et al., 1991; Schrans-Stassen et al., 1999; Yoshida et al. 2006). The cells that undergo the first wave of spermatogenesis do not pass through a Ngn3-positive, undifferentiated spermatogonia stage, but directly give rise to c-Kit expressing cells (Yoshinaga et al., 1991; Schrans-Stassen et al., 1999; Yoshida et al. 2006).

During the first wave of spermatogenesis germ cells in all tubules undergo a single synchronised wave of germ cell development (Bellve et al., 1977; Maratou et al., 2004). In contrast, to ensure a constant supply of mature sperm, in the postpupertal male germ cells undergo subsequent rounds of spermatogenesis within seminiferous tubules at different times. Spermatogonia that enter meiosis are referred to as primary spermatocytes (Maratou et al., 2004). Spermatocytes that undergo the second meiotic division are referred to as secondary spermatocytes and the resulting haploid cells are termed spermatids. The haploid spermatids then undergo spermiogenesis which involves dramatic morphological changes. In order to generate spermatozoa, spermatids extrude their cytoplasm, generate the sperm tail, and highly condense their DNA (O'Donnell et

al., 2011). During each stage of spermatogenesis cells move progressively further towards the lumen of the tubule. During spermiogenesis it is not only spermatids that undergo changes, the supporting sertoli cells that allow spermatozoa to be disengaged into the lumen of the seminiferous tubule also change and this process is called spermiation (O'Donnell et al., 2011). The spermatozoa then travel through the epididymis where they are further modified to acquire motility, and become capable of fertilization (Gatti et al., 2004).

Self-renewal of SSCs ensures that spermatogenesis takes place during the majority of a male's lifespan. The maintenance of the stem cell state and therewith self-renewal of SCC is thought to be facilitated by the adherence of SSCs to the basement membrane of the tubule (Shinohara et al., 2000). Spermatogenesis commences with the division of individual SCCs (A_{single} (A_s)) which will give rise to either two new A_s spermatogonia or a two-cell clone called type $A_{progeny}$ or A_{pr} which is connected by an intercellular cytoplasmic bridge (figure 1.12). The two-cell A_{pr} will in turn divide to form clones of 4, 8 and 16 $A_{aligned}$ cells (A_{al4-16}) leading to spermatogonial clones of increasing length (de Rooij and Mizrak, 2008). A_s, A_{pr} and A_{al} are c-Kit negative and considered undifferentiated spermatogonia (Nakagawa et al., 2007). After arrest at the G_0 stage of the cell cycle A_{al16} spermatogonia transform into A1 differentiated spermatogonia which are c-Kit positive (Schrans-Stassen et al., 1999). Subsequent mitotic divisions result in the formation of A_2, A_3 and A_4 spermatogonia (Russell et al. 1990, de Rooij & Russell 2000). A_4 spermatogonia give rise to intermediate spermatogonia which divide mitotically to produce type B spermatoginia (figure 1.13A). Type B spermatogonia divide once more and then enter the first meiotic division as primary spermatocytes (Oatley and Brinster, 2008). During cell division and differentiation spermatogonia remain connected by intercytoplasmic bridges that allow diffusion of RNAs and other solutes across all cells within a clone (Kato et al., 2004). If those bridges are not formed spermatogonia are unable to reach meiosis as demonstrated by the observation that a mutation in the above described *Tex14* gene causes male infertility (Greenbaum et al., 2006). After cells that have entered and completed meiosis further maturation takes

place, as described above. A diagram summarising the postnatal events of spermatogenesis is given in figure 1.13.

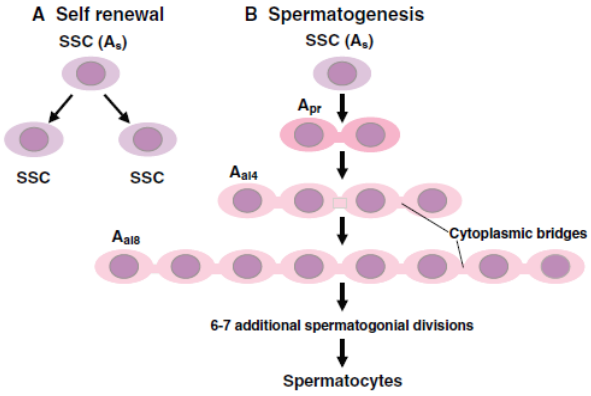


Figure 1.12 Cell fate decisions of spermatogonial stem cells. (A and B) In the mammalian testes spermatogonial stem cells (SSC) either undergo self-renewal to maintain a SSCs population that ensures a continuous supply of germ cells during the majority of the male's life span or differentiate which will result in the production of mature spermatozoa. A_s, A_{pr} and A_{al} are considered undifferentiated spermatogonia. A_s self renew by forming two new single cells or initiate differentiation by maintaining a cytoplasmic linkage between the two daughter cells (A_{pr}). The A_{pr} spermatogonia divide and form a chain of four spermatogonia (A_{al}) which divide twice more until they develop into differentiated A_1 spermatogonia. The cells undergo further cell divisions and after a total of 9-10 divisions spermatocytes form. Figure was taken from de Rooij and Mizrak, (2008).

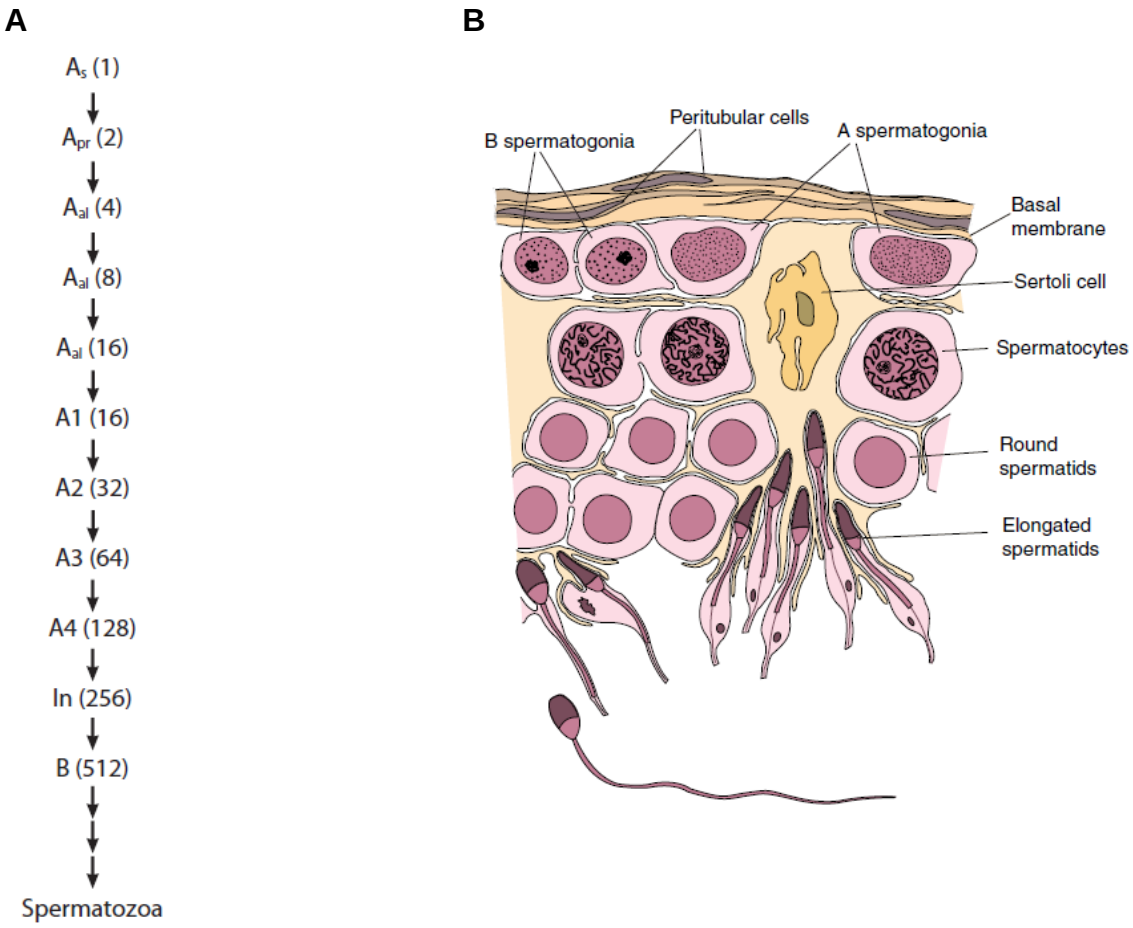


Figure 1.13: Spermatogenesis in murine testes. (A) Division of A_s can give rise to spermatogonia that are committed to differentiation (A_{pr}). A_{pr} undergo a series of mitotic divisions to produce A_{al4}, A_{al8} and A_{al16} spermatogonia. A_{al16} give rise to more mature spermatogonia, beginning with A1 spermatogonia and ending with type B spermatogonia. Type B spermatogonia enter meiosis and give rise to haploid spermatids which will undergo spermiogenesis to produce mature spermatozoa (Figure was taken from Oatley and Brinster, 2008). (B) Spermatogonia are the least differentiated germ cells in the adult testis and if committing to differentiation and subsequently enter meiosis give rise to primary spermatocytes. As spermatogenesis proceeds, germ cells (blue) progressively move towards the lumen of the tubule while retaining contact with the Sertoli cells (pink). Spermatocytes undergoing the second meiotic division are referred to as secondary spermatocytes. Meiosis results in haploid round spermatids which will undergo severe morphological changes to mature into spermatozoa. Picture was obtained from de Rooij and Mizrak, (2008).

1.6 Aneuploidy

Aneuploidy refers to an aberration from the normal chromosome compliment, euploidy, caused by gain (hyperploidy) or loss (hypoploidy) of one or more individual chromosomes (Jallepalli and Lengauer, 2001). Aneuploidy can result from errors in mitotic chromosome segregation or errors during either of the two meiotic divisions. Chromosome abnormalities of both pre- and post-zygotic origin have been postulated to be responsible for early pregnancy failures in humans (Munne et al., 1999).

1.6.1 Generation of Aneuploidy

1.6.1.1 Post-zygotic origin of aneuploidy

Aneuploidy of post-zygotic origin is usually a result of mis-segregation of chromosomes in one or few blastomeres during embryo proliferation. This results in a cytological mosaic state of cells within the developing embryo caused by an apparently random allocation of chromosomes to daughter cells (Wells et al., 1999; Munne et al., 2002; Ambartsumyan and Clark, 2008). This is also referred to as mosaic aneuploidy. Presumably the earlier that chromosome mis-segragation occurs during embryo proliferation, the greater the number of daughter cells that will inherit the abnormal karyotype. It has been suggested that the oocyte plays, at least in humans, an important role in providing RNAs and proteins required for maintaining genetic and chromosomal stability in early embryos until embryonic genome activation occurs (Braude et al., 1988). Nevertheless, little is known about the mechanisms that ensure faithful chromosome segregation during early embryonic development despite its particular importance for the survival of the embryo (Ambartsumyan and Clark, 2008). Mosaic aneuploidy also occurs in embryos with chromosomal abnormalities of pre-zygotic origin as it has been recently reported that primary aneuploid embryos progress to a mosaic aneuploid state during early development (Lightfoot et al., 2005). Recently van

Echtern-Arends and colleagues (2011) reviewed 36 studies that examined the karyotype of cells in human pre-implantation embryos and reported that 73% of 815 embryos were characterised by mosaic aneuploidy. This number is consistent with data from an earlier study (that was not included in the review) which found that 70% of all human pre-implantation embryos were idientified as having mosaic aneuploidy (Sandalinas et al., 2001). However, it is unclear if this data is representative for normal, early human development as the embryos in the studies described were derived by *in vitro* fertilisation (IVF). The fact that the gametes used to produce the embryos analysed in the studies were derived from couples that were reproductively challenged, and therefore had problems with optimal gamete function might be the reason for the high rates of mosaic aneuploidy observed in the embryos, rather than being a general feature of early human development. However, it was shown by Vanneste et al (2009) that high rates of aneuploidy are also present in early IVF embryos derived from gametes of young, healthy couples with normal fertility who were undergoing IVF because of genetic risks for hereditary diseases, rather than fertility problems. Embryos analysed just after fertilisation showed low levels of aneuploidy but at the 8-cell stage only 9% of embryos showed a completely normal chromosomeset, the majority showed abnormalities and only about 50% possessed at least some euploid blastomeres meaning that 50% of those embryos had no chromosomally normal blastomeres at all (Vanneste et al., 2009). This data suggests that the extremely high incidence of aneuploidy in pre-implantation human embryos is either an artifact of *in vitro* maturation or that early human development is extremely prone to suffer defects in chromosome segregation.

It is known that certain culture conditions can induce aneuploidy in oocytes during maturation *in vitro,* and that this might also apply for mitotic divisions of early pre-implantation embryos (Carrell et al., 2005; Requena et al., 2009). In contrast to human pre-implantation embryos, only 1-2% of embryos from fetal diagnosis have been reported to show abnormal karyotypes (Kalousek et al., 1991; Sandalinas et al., 2001; Lightfoot et al., 2005). This suggests that aneuploidy levels observed in human pre-implantation embryos represent either a consequence of *in vitro* culture that does not

occur during development of naturally conceived embryos or, given the data described above, do indeed resemble a feature of normal human pre-implantation development, that aneuploid cells or embryos, are eliminated during early development. In order to address this matter, early human pre-embryos conceived *in vivo* would need to be analysed and karyotyped. However, due to ethical reasons this does not represent a valid option and therefore it is currently not possible to clearly distinguish between the two possibilities. Comparison of aneuploidy in IVF and naturally conceived mouse embryos might provide useful information about to what extent, if at all, aneuploidy rates differ between *in vitro* and *in vivo* developing embryos.

1.6.1.2 Pre-zygotic origin of aneuploidy

In addition to aneuploidies of post-zygotic origin aneuploidies of pre-zygotic origin represent a major cause of the formation of aneuploid embryos as aneuploid gametes will transmit the abnormal karyotype to their offspring if fertilised. Aneuploidy transmitted to offspring through the parental germline may be caused by mis-segregation of chromosomes during either of the two meiotic divisions. Failure to resolve the physical connection or chiasmata between homologous chromosomes during the first meiotic division results in both homologues segregating together to the same pole which is referred to as "true" non-disjunction (Hassold and Hunt, 2001). Homologous chromosomes segregating to the same pole can also originate from premature resolution of chiasmata or failure to establish crossovers in the first place. This defect, referred to as "achiasmate" non-disjunction, might result in homologues segregating independently to the same pole (figure 1.14) (Hassold and Hunt, 2001). Furthermore, premature sister chromatid separation might result in the segregation of a whole homologue and a single chromatid to each spindle pole (figure 1.15) (Hassold and Hunt, 2001). Aneuploidy resulting from the second meiotic division is generally caused by a failure of sister chromatid separation (figure 1.14). However, which of these constitutes the major cause of aneuploidy observed in the human female germline is still subject to controversy. Studies from human trisomies suggest that maternal MI errors are likely an important

contributor to most human aneuploidies (Wolstenholme and Angell, 2000; Rosenbusch, 2004; Pellestor et al., 2005; Hassold et al., 2007). For example, a comparison of 1070 cases of trisomy 21 provides evidence that the majority of cases are caused by maternal MI errors associated with failure in recombination or distally located chiasmata (Allen et al., 2009). Trisomy 21 cases scored as arising at maternal MII have been associated with extremely proximal chiasmata suggesting that they in fact derive from recombination errors that affect MI segregation (Risch et al., 1986). Both maternal MI and so-called MII errors increase with maternal age but interestingly the recombination errors underlying trisomy 21 differ between age groups (Oliver et al., 2009). In women under 30 years of age failure of recombination is the most frequent cause of trisomy 21 whereas this particular defect is not prevalent in women between 30-34 years. In women older than 35 trisomy 21 caused by recombination failure is elevated again (Hassold and Hunt, 2009; Oliver et al., 2009). Furthermore trisomy 21 caused by distally located chiasmata appears to decline with increasing maternal age whereas trisomy 21 caused by pericentromeric exchanges in MII errors seems to increase (Oliver et al., 2009). It is not yet clear what these age-related recombination changes mean but different recombination patterns have also been linked with the fidelity of chromosome segregation for chromosomes other than chromosome 21 during meiosis (Hassold et al., 2007; Hassold and Hunt, 2009). Those studies suggest that homologues chromosomes with no crossovers or crossovers too far (distal) or too close (proximal) to the centromeres are prone to non-disjunction in humans. For example, in addition to chromosome 21, trisomies 16 and sex chromosomes trisomies have been correlated with alterations in the location of cross-over formation whereas reductions in recombination appear to predispose chromosomes 13, 15, 16, 18, 21 and 22 as well as the sex chromosomes to mis-segregate (Bugge et al., 1998; Robinson et al., 1998; Thomas et al., 2001; Lamb et al., 2005; Sherman et al., 2006; Bugge et al., 2007; Hall et al., 2007a, Hall et al., 2007b). The observations from those studies suggest that sub-optimal recombination events occurring in the fetal oocyte contribute to human non-disjunction. Other groups have postulated that premature sister chromatid separation, rather than non-disjunction of whole bivalents at meiosis I, represents the main source of

chromosomal aberrations in our species (Angell, 1991; Angell et al., 1991; Pellestor et al., 2002). In accordance with that Vilard et al. (2006) found that among 43 abnormal human oocytes out of 141 oocytes 80% of cases showed premature sister chromatid separation. Furthermore, premature sister chromatin separation has been proposed to be the main underlying mechanism of age-related aneuploidy in humans (Angell et al., 1994; Angell, 1997; Wolstenholme and Angell, 2000). Taken together those data suggest that there are likely multiple causes of human age-related nondisjunction and that some chromosomes might be more susceptible to certain defects than others. This is supported by a study by Garcia-Cruz et al (2010) where the authors show that among 46 human oocytes premature sister chromatid separation events are most prevalent for chromosome 22, followed by chromosome 15 whereas vulnerable patterns of meiotic recombination appear characteristic for chromosome 16.

It has been proposed that the maternal age related effect of human aneuploidy is the consequence of a two-hit failure process (Orr-Weaver, 1996). This model suggests that the first hit is caused by nondisjunction-prone chiasmata configurations in the human fetal oocyte and the second hit results from an age-dependent component which reduces the ability of the cell to partition the error prone chromosomes (Lynn et al., 2005). Different mechanisms for the second hit have been proposed including degeneration of centromeric cohesin, spindle aberrations and chromosome misalignment, depletion of oocyte pools and a deterioration of the SAC with increasing maternal age (LeMaire-Adkins et al., 1997; Volarcik et al., 1998; Kline et al., 2000; Hodges et al., 2005; Cheng et al., 2009; Chiang et al., 2010; Lister et al., 2010; Chiang et al., 2011).

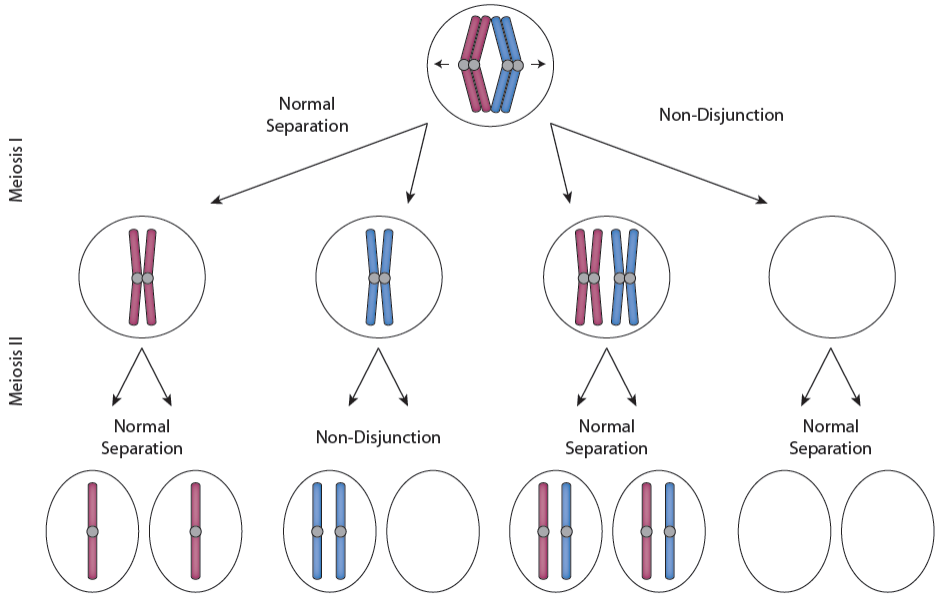


Figure 1.14 Achiasmate non-disjunction during meiosis. A normal meiosis I (MI) achieves the segregation of homologue chromosomes (purple and blue). Mis-segregation can occur when homologues fail to separate and travel together to the same pole or when homologues that have failed to pair and/or recombine travel independently to the same pole. Normal meiosis II facilitates the segregation of sister chromatids. Non-disjunction at MII is assumed to result from failure of the sisters to separate.

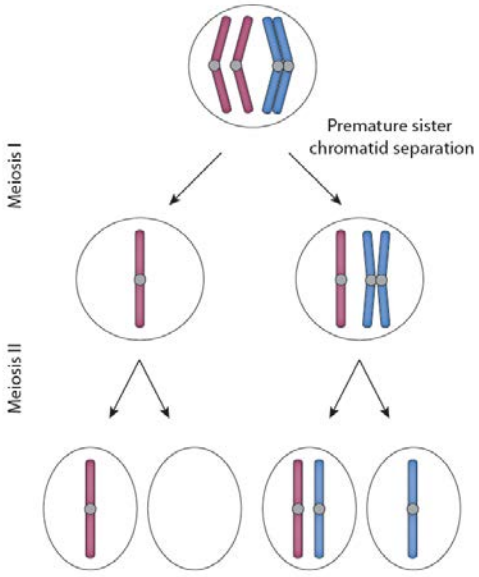


Figure 1.15 Premature sister chromatid separation Mis-segregation of homologue chromosomes (purple and blue) during MI can result from premature separation of sister chromatids (purple). In this case chromatids, rather than homologues, segregate from one another. It has been reported that in addition to mutations affecting the cohesin complex pericentromeric exchanges might disrupt sister chromatid cohesion, resulting in the premature separation of sisters at MI (Hassold and Hunt, 2001).

1.6.2 Consequences of aneuploidy

Aneuploidy generally results in inviability. Therefore it is not surprising that in most organisms the likelihood of chromosome mis-segregation during meiosis is very rare. For example in the yeast *Saccharomyces cerevisiae* the frequency of aneuploidy is estimated to be as low as 1 in 10,000 and also only around 1 in 3,000 in the fruit fly *Drosophila melanogaster* (Sears et al., 1992; Koehler et al., 1996). In the mouse the incidence of meiotic errors seems to be higher, however, the overall incidence of aneuploidy among zygotes does not exceed 1–2% (Hasshold and Hunt, 2001). In contrast an estimated 10-30% of all fertilized human oocytes are aneuploid. Aneuploidy is the most common chromosome abnormality of our species and the dosage imbalance of chromosomes is a major cause of reproductive failure (Hasshold, 2007). As demonstrated by many mouse models defects in meiotic chromosome behaviour during spermatogenesis produce aneuploid gametes leading to spermatocyte cell death that results in male sub- or infertility. For example several meiotic mutations affecting processes that take place during prophase of meiosis, such as mutations in genes that are involved in early recombination or synaptonemal complex formation, trigger meiotic arrest and subsequent death of spermatocytes during mid-pachytene, resulting in reduced fertility or infertility (Yuan et al., 2000; Yang et al., 2006; Bolcun-Filas et al., 2007; Hamer et al., 2008). In contrast, in females, germ cells respond less strictly to meiotic defects and oocytes tend to tolerate some states of aneuploidy (LeMaire-Adkins et al., 1997; Hunt and Hassold, 2002; Yuan et al., 2002; Kouznetsova et al., 2007). Females with mutations in certain meiotic genes are fertile but pass aneuploidies on to their progeny leading to high rates of dysmorphology and mortality within the first days of embryonic development (Eichenlaub-Ritter, 2005).

Approximately 7–10% of clinically recognized pregnancies are known to be chromosomally abnormal (Hassold and Hunt, 2001). However, the total numbers of aneuploid pregnancies is likely to be underestimated because the loss or gain of some human chromosomes is lethal and therefore these pregnancies may not survive long

enough to reach clinical attention. It is believed that aneuploidies constitute the leading cause of miscarriage in humans (Wells and Delhanty, 2000; Pacchierotti et al., 2007). In cases where the aneuploidies are compatible with live birth they are responsible for various clinical syndromes. Aneuploidy caused by loss or gain of one chromosome are also referred to as monosomy or trisomy, respectively. The only known monosomy which is able to survive until birth is Turner syndrome that is caused by loss of one X chromosome. In contrast, several trisomies survive until birth. Among those are sex chromosome trisomies like Klinefelter syndrome (XXY) and triple X syndrome (XXX). It is possible that sex chromosome aneuploidies are reasonably well tolerated, because as in a normal situation, all but one of the X chromosomes get inactivated and thus the X-linked gene dosage remains one (Chow et al., 2005). Although about 15% of X-linked genes escape inactivation the over-expression of those appears to be compatible with life (Carrel et al., 2005). The Y chromosome itself comprises relatively few protein-coding genes, 78 in total, and increased dosage from the Y chromosome seems compatible with life (DesGroseilliers et al., 2002; Skaletsky et al., 2003). Whole chromosome aneuploidies of most autosomes are not viable and those trisomies that do survive after birth can have severe phenotypes like trisomies of chromosomes 13, 18 and 21 (Hassold and Hunt, 2001). Trisomy 13 (Patau syndrome) and trisomy 18 (Edwards syndrome) are hallmarked by a very short life expectancy (days and months, respectively), mental retardation as well as multiple developmental abnormalities. Trisomy 21 or Down's syndrome, the paradigm of whole chromosome aberrations in human development, is the mildest of all the autosomal aneuploidies and is compatible with a relatively long life (Roper and Reeves, 2006). The reason for trisomy 21 being the mildest is likely due to the fact that chromosome 21 is the smallest autosome with the fewest genes. However, it too results in a range of developmental abnormalities and is the most common known genetic cause of mental retardation (Antonarakis et al., 2004).

Several studies in mice showed that aneuploid mouse embryos are, despite having an abnormal karytope, and depending on the particular chromosome gained or lost, often capable of cell division, differentiation as well as implantation and initiation of the

gastrulation process. However, most aneuploid embryos are not able to sustain gastrulation and/or suffer growth and developmental retardation (Hernandez and Fisher, 1999; Lightfoot et al., 2005). The defects arising at gastrulation might be a result from aneuploid cells undergoing apoptosis at this stage and this is discussed below. Robertsonian chromosomes are formed by joining single chromosomes at the centromere ends. This has no immediate effect on the phenotype of the organism but confers the carrier with a higher risk of chromosome mis-segregation during meiosis. By taking advantage of mouse strains carrying Robertsonian chromosomes it was shown that severity of developmental defects that occur in aneuploid embryos depends on the type of aneuploidy and on the particular chromosome gained or lost (Hogan et al., 1994). Aneuploid mouse embryos can be produced selectively for any chromosome by crossing mice doubly heterozygous for two different Robertsonian chromosomes involving the same chromosome (Hogan et al., 1994). Studies using this approach showed that mice suffering from autosomal trisomies or monosomies were characterised by variable phenotypes depending on which chromosome is affected. Autosomal trisomies of chromosomes 2, 7, 8 or 15 are embryonic lethal before 10 days gestation whereas chromosome 3, 4, 5, 9, 11 or 17 trisomies are lethal by 10-13dpc and 1, 6 and 10 by 15dpc (Gropp et al, 1975; Gropp et al 1983; Gearhart, J. et al, 1986; Dyban and Baranov, 1987; Gearhart, et al, 1987; Beechey and Searle 1988; Smith and Walker, 1992; Cacheiro et al 1994). Only very few trisomic embryos survive until birth (chromosomes 12, 13, 14 or 18) and only embryos with trisomy 19 have been reported to survive beyond birth (Gropp et al 1983; Gearhart et al, 1986; Morriss-Kay and Putz 1986; Dyban and Baranov, 1987; Ninomiya et al, 1993; Lorke, 1994). Monosomies are even more detrimental than trisomies and in the case of most chromosomes results in embryonic lethality at the blastocyst stage or earlier (Kaufman and Sachs, 1975; Epstein and Travis, 1979). Observations regarding the fate of embryos from *Sycp3*$^{-/-}$ females suggest that aneuploid cells undergo apoptosis in the post-implantation epiblast. *Sycp3*$^{-/-}$ females transmit aneuploidy to about 1/3 of their offspring and it has been shown that those embryos exhibit morphological abnormalities by 7.5dpc and embryo structure is lost by 8.5dpc. This suggests that embryos with abnormal karyotype succumb early in

gestation during gastrulation. It has been shown that this is mediated through a p53-independent apoptotic mechanism (Lightfoot et al., 2005).

As mentioned above, the majority of aneuploidies are of maternal origin resulting from errors during the first meiotic division (Hassold and Hunt, 2001). Between 1% and 4% of sperm and as many as 10-30% of human oocytes have been estimated, by molecular cytogenetic analysis, to be aneuploid (Hasshold and Hunt, 2001; Pacchierotti et al., 2007). In addition to the gender effects on the incidence of aneuploidy, there exists a striking relationship between maternal age and aneuploidy. The risk for aneuploidy among women in their early 20s in a clinically recognized pregnancy is 2-3% whereas among women which are at the end of their reproductive lifespan the risk dramatically increases to at least 50% (Hassold and Chiu, 1985). The basis for the effect of maternal age on the occurrence of aneuploidy or for the general difference in incidence of aneuploidy between our own and other species is still unclear. Two recent studies show that mammalian female aging, and the occurrence of chromosome mis-segregation, correlate with decreased levels of the cohesin protein *Rec8* and the cohesion protector protein *Sgo2* in mice. This suggests that age related loss of centromere cohesion in older females might contribute to the increased levels of aneuploidy in aging women (Chiang et al., 2010; Lister et al., 2010). However, given the complexity of the process of chromosome segregation, and in particular meiosis, it is likely that several problems contribute to the high frequency of female transmitted aneuploidy observed in humans. Meiotic mouse mutants provide valuable tools to model the gain or loss of whole chromosomes during meiosis in order to understand the mechanisms that underlie human aneuploidy (Hassold et al, 2007). Existing mouse models for aneuploidy are summarised below.

1.6.3 Existing mouse models for aneuploidy

The study of human female meiosis is challenging and complicated given the long life cycle of the oocyte and the technical, as well as ethical difficulties, to obtain and study material from humans. As reliable protocols for the induction of mammalian meiosis *in vitro* currently do not exist, mice represent an important model system to advance our understanding about the mechanisms that may cause or contribute to chromosome mis-segregation.

The characterisation of a large number of mouse knockout models has linked mutations in genes involved in synapsis, recombination, sister chromatid cohesion, and other meiotic pathways involved with chromosome mis-segregation. These models also showed that female germ cells appear to respond less strictly than male germ cells to certain abnormalities that can arise during meiotic chromosome synapsis and segregation. This difference is despite the fact that both male and female germ cells possess meiotic checkpoint mechanisms that detect abnormal chromosome behavior during meiosis in order to prevent the formation of aneuploid gametes (Morelli and Cohen, 2005; Jones, 2008). Male and female mice that lack genes functioning during meiotic recombination such as *Spo11*, *Dmc1*, *Msh4* and *Mlh1* are typically infertile. *Spo11*, *Dmc1* and *Msh4* mutants show defects in paring and synapsis of homologue chromosomes and germ cells undergo arrest at prophase I (Pittman et al., 1998; Baudat et al., 2000; Kneitz et al., 2000; Romanienko and Camerini-Otero, 2000). Mutations in *Mlh1* also cause sterility in both sexes but pairing and synapsis of homologous chromosomes is not disrupted. However, recombination sites are unable to mature, leading to the formation of achiasmate meiotic chromosomes that result in the majority of chromosomes being found as univalents at pro-metaphase which will trigger meiotic arrest and apoptosis. Similarly mutations in genes involved in chromosome synapsis cause defects in meiotic progression. Mice with mutations in the lateral elements *Sycp2* or *Sycp3* of the synaptonemal complex are characterised by impaired chromosome synapsis as mentioned under 1.4 (Yuan et al., 2002; Yang et al., 2006). In males this

triggers apoptosis at the pachytene stage of meiosis resulting in infertility. In contrast *Sycp2*$^{-/-}$ and *Syc3*$^{-/-}$ females are subfertile. It has been shown that subfertility, at least in *Sycp3*$^{-/-}$ females, is caused by a failure in bivalent formation resulting in high rates of aneuploidy in the mature oocytes. It appears that differences in the phenotype of *Sycp3*$^{-/-}$ females which are subfertile and *Mlh1*$^{-/-}$ which are infertile, albeit both being characterised by formation in achiasmate meiotic chromosomes, is caused by different numbers of univalent chromosomes present at pro-metaphase I. Kouznetsova and colleagues observed that a small number of univalents, as observed in *Sycp3*$^{-/-}$ females, avoids detection at pro-metphase resulting in progression of meiosis, whereas in *Mlh1*$^{-/-}$ females the large number of chromosomes found as univalents will cause meiotic arrest (Edelmann et al., 1996; Kouznetsova et al., 2007). Mutations in the meiosis specific cohesin subunit *Smc1β* cause aberrant chiasmata formation in oocytes with crossovers frequently being located closer to the telomeres and connection between homologous chromosomes frequently lost by pro-metaphase I (Hodges et al., 2005). This effect appeared exacerbated with maternal age as in addition to loss of connection between homologous chromosomes, older females also show premature loss of sister chromatid cohesin. Similarly mutations in *Rec8* cause impaired synapsis and therefore problems with recombination between homologous chromosomes leading to sterility in both sexes. These models elucidate that all stages of meiotic pro-phase are crucial and must occur properly for faithful chromosome segregation to take place.

It has been suggested that in humans sequence divergence between chromosomes causes reduced meiotic recombination resulting in increased achiasmate non-disjunction (Borts et al., 2000). In 1993 Hale et al. showed that F1 female progeny from C57BL/6J x *Mus spretus* crosses, two inbred mouse strains with about 1% sequence divergence, suffered defects in chromosome synapsis and reduced fertility. A later study by Koehler et al. (2006) analysed the oocytes of the F1 progeny from C57BL/6J x *Mus spretus* crosses in more detail and observed, in addition to defects in homologous chromosome synapsis, disturbed meiotic recombination and increased levels of achiasmate non-disjunction. The level of aneuploidy in 4 weeks old F1 female progeny was about 10% (1% in

controls) and increased to 20% (2.6% in controls) in 8-11 month old females (Koehler et al., 2006). The increase in aneuploidy in older females in this study might be explained by the centromeric loss of cohesin observed in older female mice as described by Chiang and colleagues (2010) and Lister and colleagues (2010).

1.7 Germ cells, pluripotent stem cells and retrotransposons

In the last 20 years, progress has been made towards understanding the occurrence of aneuploidy in humans. However, it remains largely obscure why humans show much higher rates of aneuploidy than other species, or why the risk of transmitting aneuploidies increases with maternal age in humans. In recent years a number of mouse mutants have been generated that have defects in silencing retrotransposons. Interestingly, in many of these models the de-repression of retrotransposons correlates with defects in meiotic progression and the occurrence of aneuploidy (Kuramochi-Miyagawa et al., 2004; De La Fuente et al., 2006; Carmell et al., 2007; Öllinger et al., 2008, Soper et al., 2008).

1.7.1 Retrotransposons

Transposon elements are stretches of DNA that can move and multiply within the genome of an organism and were first described by Barbara McClintock during the 1940s (McClintock, 1956). Transposons must be active in the germline, or in pluripotent cells in early embryos that can give rise to germ cells, in order to amplify themselves in the host genome through evolutionary time. Genome-sequencing projects have shown that the mammalian genomes contain large numbers of transposon elements which have accumulated over millions of years by being copied into new genomic locations (Giordano at al., 2007). Type I transposon elements or retrotransposons, constitute the major class of transposon elements in mammals and comprise around 37.5% of the

mouse genome (Mouse genome sequencing consortium). Retrotransposons are mobile genetic elements that amplify and move themselves to new locations in the genome. Retrotransposons can be divided into two classes on the basis of either the presence or absence of long terminal repeats (LTRs). Non-LTR retroelements are the autonomous long interspersed elements (*LINEs*) and the non-autonomous short interspersed elements (*SINEs*). *LINE-1* encodes two proteins required to mediate retrotransposition and these proteins are also used by *SINEs* to retrotransposons as *SINEs* are derived from endogenous small cellular RNAs that lack protein-coding capacity of their own (Martin, 1991; Wallace et al., 2008). The entire class of non-LTR retrotransposons is at least 600 million years old (Malik et al., 1999*)*. Over evolutionary time non-LTR retrotransposons have increased the diversity of the genome through a variety of mechanisms and at present continue to sculpt mammalian genomes by behaving as insertional mutagens. The LTR containing retroelements are the endogenous retroviruses (ERVs) and are found in the genomes of numerous eukaryotes and have similar structures to simple retroviruses. An exogenous retrovirus (XRVs) occasionally may infect the germline which can lead to an integrated provirus that is passed to the offspring and inherited in a mendelian fashion and which will then be called an ERV (Stoye, 2001). The common feature of non-LTR and LTR retrotransposons is that they reproduce through an RNA intermediate which is reverse-transcribed back into DNA by a reverse transcriptase to allow integration back into the genome (figure 1.16). Whenever a retrotransposon lands in the host genome it may destabilize the genome by modifying gene structure, altering or abolishing gene expression, by deleting large blocks of DNA or by favouring rearrangements through recombination between non-allelic repeats (Yoder et al., 1997).

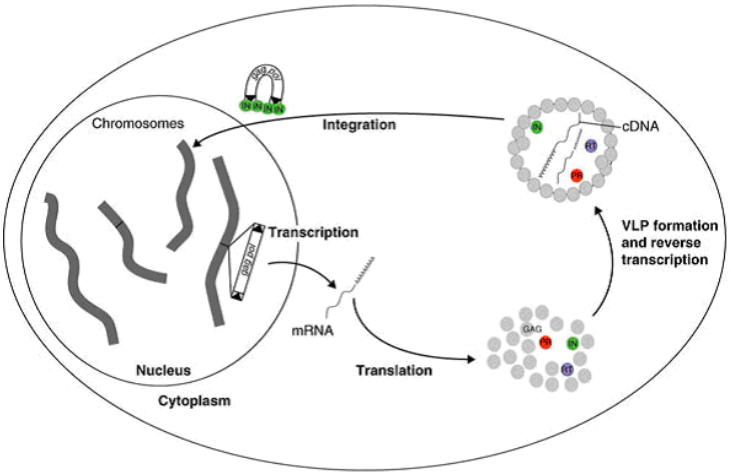


Figure 1.16 The life cycle of LTR retrotransposons The retrotransposon life cycle starts with transcription of the in the host genome residing, element. The region that encodes the viral structural and enzymatic proteins is flanked by LTRs. Transcription is followed by translation of the produced RNA resulting in protein expression of genes carried by the virus. The *gag* gene encodes structural proteins that form the virus-like particle (VLP) where reverse transcription will take place. The *pol* gene encodes several enzymatic functions, including a protease that cleaves the Pol polyprotein, a reverse transcriptase (RT) that copies the retrotransposon's RNA into cDNA, and an integrase (IN) that integrates the cDNA back into the genome (taken from Havecker et al., 2004).

1.7.2 Mechanisms to control transposon elements in pluripotent stem cells and germ cells

It is believed that retrotransposons and other transposable elements have shaped the evolution of genomes in which they reside for tens or even hundreds of millions of years by generating insertion mutations or altering gene expression by functioning as regulatory units for host genes and contributing to genetic innovation (Speek, 2001; Jurka et al., 2007). Furthermore, it has been suggested by Peaston et al. (2007) that retrotransposons might act as alternative promoters and/or as first exons for certain genes in oocytes and cleavage stage embryos and therefore regulate expression of those genes during early development. This and the generation of genetic diversity through *de novo* transposition events might, at least in part, have contributed to the evolutionary success of retrotransposons which is reflected by the fact that retrotransposons comprise a major part of mammalian genomes. However, mutations and deletions caused by

retrotransposons will, in many cases, be harmful and correlate negatively with the fitness of the host. Therefore, germ cells and pluripotent cells may need to limit the level of retrotransposon activity. It has been shown that in mice the developing germ cells as well as pluripotent cells appear to possess mechanisms to silence expression of retrotransposons and limit their mutagenic activity (described below) but it is unknown if this is the case in other species. Mutations in genes involved retrotransposon repression result in increased expression of retrotransposons and presumably defects during embryonic development and meiosis.

1.7.2.1 Retrotransposon silencing in embryonic stem cells

As mentioned above, retrotransposons have to reproduce in germ cells or pluripotent cells in the early embryo, which can give rise to germ cells, in order to manifest themselves in the host genome through the generations. One of the major mechanisms to restrict expression of retrotransposons used by somatic, pluripotent as well as germ cells is transcriptional repression. DNA methylation plays a role in transcriptional repression of retrotransposons in somatic cells and in germ and pluripotent cells. The majority of DNA methylation in the mammalian genome is thought to be associated with retrotransposon sequences and some of those are activated in the event of genome wide demethylation (Yoder et al., 1997). However, pluripotent cells appear to depend less on this mechanism as ES cells with mutations in all three catalytically active DNA methyltransferases (*Dnmt1*$^{-/-}$*Dnmt3a*$^{-/-}$ *Dnmt3b*$^{-/-}$ triple knock out (*Dnmt* TKO) ES cells) show modest upregulation, around ~4-8 fold, of *IAP*, *RLTR45*, *RLTR1B* and *MMERGLN* elements (Reichmann et al., submitted). This level of upregulation is small compared to the upregulation of *IAP* elements seen in differentiated somatic tissues in response to hypomethylation as *Dnmt1*$^{-/-}$ fibroblasts upregulate *IAP* expression by about 50 to 100-fold (Jackson-Grusby et al., 2001). The is caused, at least in part, by the fact that ES cells rely on the transcriptional co-repressor Kap1 to repress IAP elements (Hutnick et al., 2010; Rowe et al., 2010; Karimi et al., 2011). *Kap1* presumably acts by recruiting the histone H3K9 methyltransferases *Eset/Setdb1* to *IAP* as well as *MERVL*, *MusD* and

MuLV retrotransposon chromatin loci. However, not all retrotransposons are targeted by *Kap1* and/or show enrichment for the repressive histone modification H3K9 methylation and therefore it is likely that additional mechanisms play a role in transcriptionally repressing these retrotransposons in ES cells. It has also been shown that pluripotent cells restrict retrotransposon expression through a variety of mechanisms. For example the Dicer mediated siRNA pathway has been reported to regulate *IAP* and *MuERVL-L* endogenous retrovirus transcript abundance in mouse pre-implantation embryos as well as *LINE1* and *IAP* retrotransposon expression in ES cells (Calabrese et al., 2007; Kanellopoulou et al., 2005; Hutnick et al., 2010). However, these changes appear to be rather subtle, around 2-fold, which suggests that this pathway is probably not a major mechanism to silence retrotransposon expression in ES cells. A recent study by Kano et al. (2009) suggests that *LINE1* RNA produced in germ cells is passed on to the zygote where it can get reverse transcribed and integrated in the genome during the early cleavage stages. This implies that pluripotent cells might have further post-transcriptional mechanisms in place in addition to Dicer-dependent RNAi in order to restrict the retrotransposition of inherited retrotransposon transcripts. It also suggests that germ cells are not only required to repress retrotransposon activity to protect their own genome but also to avoid retrotransposon RNA that is carried over by either oocyte or sperm into the next generation.

1.7.2.2 Retrotransposon silencing in male and female germ cells

As mentioned above both pluripotent cells and especially somatic cells rely on DNA methylation to silence certain retrotransposons. However, DNA methylation patterns in the developing germline are dynamic and differ between the sexes. PGCs undergo, as described under 1.3, widespread DNA demethylation soon after colonising the gonad when genetic imprints are erased. Accordingly, this provides a developmental window where transposon elements are not silenced by DNA methylation. PGCs show partial demethylation of IAP elements by 11.5dpc and over the following few days, DNA methylation of retrotransposons is further reduced (Hajkova et al., 2002; Lees-Murdock

et al., 2003). Subsequent de novo methylation sequential to the genome wide hypomethylation has a major impact on fertility by resetting gametic imprinted methylation marks, according to the sex of the individual, and by resilencing of transposon elements.

In the male germline *de novo* methylation of retrotransposons starts around 14.5dpc and *IAP* elements are fully methylated at 17.5dpc, well in advance of meiosis and this is believed to be constantly maintained throughout the adult life of the organism (Lees-Murdock et al., 2003). In male germ cells *Dnmt3L*, which acts as a co-factor for the de novo methyltransferases, and two of the mouse piwi-like genes, *Mili* and *Miwi2* are required to establish de novo methylation of retrotransposons in mitotically quiescent prospermatogonia (Kuramochi-Miyagawa et al., 2008). *Mili* as well as *Miwi2* contain a piwi domain that is implicated in binding and cleaving RNA molecules (piRNAs) (Kuramochi-Miyagawa et al., 2004; Carmell et al., 2007). It is thought that piRNAs directly influence *de novo* methylation by acting as a sequence specific guide for the methylation machinery. However, a more indirect relationship between piRNA/Piwi-like protein complexes and *de novo* DNA methylation cannot be excluded at present (Unhavaithaya et al., 2009). Consistent with the idea that piRNA/Piwi-like protein complexes influence *de novo* methylation, *Dnmt3L*$^{-/-}$, *Miwi2*$^{-/-}$ and *Mili*$^{-/-}$ male germ cells show remarkably similar phenotypes. Specially, the mutants are characterised by loss of methylation and subsequent upregulation of *IAP* and *LINE1* elements in neonatal testis which correlates with defects in chromosome synapsis resulting in germ cell apoptosis during pachytene and therefore infertility. Several other genes have been identified to function in the piRNA pathway and are also characterised by retrotransposon derepression and meiotic defects in the male germline (Chuma et al., 2006; Soper et al., 2008; Ma et al., 2009; Reuter et al., 2009; Shoji et al., 2009; Frost et al., 2010; Zheng et al., 2010).

In contrast the function of the piRNA pathway in female germ cell remains yet to be elucidated. *Miwi2* expression is mainly restricted to the male germ line but *Mili* is also

expressed in female germ cells. *Mili*$^{-/-}$ females are fertile but show about a 3.5-fold increase in *IAP* RNA abundance in growing oocytes (Kuramochi-Miyagawa et al., 2001; Watanabe et al., 2008). This suggests that *Mili* is involved in retrotransposon silencing in the female germ line but that the de-repression of *IAP* transcripts that occurs in the *Mili*$^{-/-}$ oocytes is not sufficient to affect fertility. It will be interesting to determine how exactly *Mili* silences retrotransposons in the female germline. *De novo* methylation of retrotransposons takes also place in the female germline but later than in male germ cells during oocyte growth after the oocytes have progressed through the early meiotic prophase. Furthermore, in females *de novo* methylation of retrotransposons appears not to be *Dnmt3L* dependent. Still, DNA methylation has been proposed to be required for silencing retrotransposons and progression through meiosis in the female germline (De La Fuente et al., 2006). This is based on data from *Lsh*$^{-/-}$ female embryos which show reduced DNA methylation at *IAP* retrotransposons in meiotic oocytes and that these oocytes fail to progress through early meiotic prophase due to severe impairment of homologous chromosome synapsis (De La Fuente et al., 2006). The meiotic phenotype of *Lsh*$^{-/-}$ males has not been determined as *Lsh*$^{-/-}$ mice die perinatally or a few days after birth before prospermatogonia enter meiosis (Geiman et al., 2001; Sun et al., 2004). Similar to pluripotent cells, female germ cells use siRNA mechanisms to constrain expression of transposable elements. Mouse oocytes with a tissue-restricted lack of the Dicer mediated siRNA pathway are not able to progress through meiosis I. Specifically, oocytes that lack Dicer show impaired spindle formation with misaligned chromosomes and upregulation of *SINE* elements and the *RLRT10*, *MT* and *MTA* endogenous retroviruses (Watanabe et al., 2008; Murchison et al., 2007). However, since Dicer also seems to be involved in the regulation and turnover of maternal transcripts it is not possible to clearly associate the retrotransposon upregulation with the observed phenotype (Murchison et al., 2007).

Given the huge number of different retrotransposon families in the mouse genome, it is likely that germ cells and pluripotent cells use multiple mechanisms to silence

retrotransposons, and that different retrotransposon families have varying susceptibilities to different silencing mechanisms.

1.8 Tex19.1

In addition to the mechanisms described above our group found a role for *Tex19.1* in the suppression of endogenous retroviruses. *Tex19.1* is a mammalian-specific gene of unknown biochemical function that is expressed in germ cells and pluripotent stem cells. It was originally isolated as a 'Testis-expressed' gene in a screen to identify genes expressed in spermatogonia but not somatic cells and was also found to be a putative target of the germline specific RNA-binding protein Dazl by Howard Cooke's group (Page et al., 2001; Reynolds et al., 2005; Childs, 2006; Kuntz et al., 2008; Öllinger et al., 2008). *Tex19.1* is well conserved among mammals but does not share sequence similarities with any known protein domains which makes speculations about its function difficult (Kuntz et al., 2008). Humans have a single *TEX19* gene with a premature stop codon in place of a conserved tryptophan residue that truncates the protein to 164 amino acids in length if used (Childs, 2006). The codon is conserved in macaque and chimpanzee and in all three species the subsequent 111 residues are conserved after the premature stop codon (Childs, 2006). In humans there is an additional premature stop codon, thirty-two residues after the first one (Childs, 2006). The conservation of the first stop codon in human, chimp and macaque indicates that the primate lineage has coped with a single, substantially-truncated gene for at least 23.3 million years since the last common ancestor (Kumar et al., 2005). Why the open reading frame downstream of the stop codon has failed to degenerate over such a long time is open to speculation. It is not clear at this point if the mRNA is translated into protein and if the truncated version would be functional. It is possible that the mRNA is a substrate for mRNA editing or other events that might facilitate a read through the stop codon/s resulting in a full length version of the protein. This would be in accordance with the fact that the amino acid sequence downstream of the stop codon has been conserved, suggesting it is still under selective pressure. However, the NCBI expressed

sequence tags database (ESTdb) does not provide evidence for this mechanism. In contrast to the primate lineage there are two *Tex19* genes in the mouse that have been generated in rodents by a duplication event that gave rise to *Tex19.1* and *Tex19.2*. Tex19.1 is a cytoplasmic protein of 354 amino acids in length. It is expressed in the pre-implantation embryo as well as pluripotent cells like embryonic stem cells (ES) and embryonic germ cells (EG) and its expression is switched off upon ES cell differentiation and in the somatic cells of adult mice (Childs, 2006; Kuntz et al., 2008). In adult mice, Tex19.1 expression is restricted to the gonads (Childs, 2006; Kuntz et al., 2008). This expression pattern suggests that *Tex19.1* could have a role in pluripotency or germline function. In order to analyze the potential role of *Tex19.1* in germ cell and pluripotent cell function, *Tex19.1*$^{-/-}$ knockout mice have been generated in the lab and the role of *Tex19.1* in the male germline has been characterized (Öllinger et al., 2008)

1.8.1 Retrotransposon upregulation and fertility defects in *Tex19.1*$^{-/-}$ males

In the male gonad *Tex19.1* is expressed in the cytoplasm of spermatogonia and early spermatocytes but not in further differentiated late pachytene spermatocytes (Öllinger et al., 2008). Deletion of *Tex19.1* results in impaired spermatogenesis (Öllinger et al., 2008). *Tex19.1*$^{-/-}$ knockout spermatocytes initiate meiotic recombination and assemble the AEs of the SC but synapsis of homologues chromosomes is impaired in about half of these cells (Öllinger et al., 2008). In addition, two thirds of metaphase I spermatocytes exhibit univalent chromosomes in *Tex19.1*$^{-/-}$ animals compared to around 6% in controls (Öllinger et al., 2008). Univalent chromosomes refer to a state where the homologues chromosomes, that should be paired up as bivalents at metaphase I, have prematurely separated.

Currently, it is not clear whether the univalents in the *Tex19.1*$^{-/-}$ spermatozytes might have arisen from the asynapsed chromosomes observed during pachytene or, more likely, were generated from normal synapsed pachytene cells. The latter would suggest a role for *Tex19.1* after pachytene in establishing or maintaining the bivalent. However,

similar to the other mouse mutants discussed above the asynapsis observed in the *Tex19.1*$^{-/-}$ males correlate with an increase in abundance of retrotransposon transcripts. *Tex19.1*$^{-/-}$ knockout males show upregulation of RNA from the *MMERVK10C* endogenous retrovirus in the testes and this transcript accumulates specifically in meiotic spermatocytes but not in mitotic spermatogonia in the absence of Tex19.1 (Öllinger et al., 2008). This retrotransposon upregulation occurring specifically in meiotic cells differs from previously described mutants since *Dnmt3L*, *Miwi* and *Mili*$^{-/-}$ males show de-repression also in the mitotically dividing spermatogonia. *Tex19.1*$^{-/-}$ do not upregulate transcripts of IAP or LINE1 retrotransposons and hence element specificity as well as the timing of retrotransposon de-repression differ between *Tex19.1*$^{-/-}$ mice and *Dnmt3L*$^{-/-}$, *Miwi2*$^{-/-}$ and *Mili*$^{-/-}$ mice. There is also no detectable change in the DNA methylation status of *MMERVK10C* retrotransposons in *Tex19.1*$^{-/-}$ knockout testes which suggests that *Tex19.1* does not appear to be involved in the *Dnmt3L*/*Mili*/*Miwi2*-dependent mechanism that methylates and silences retrotransposons in quiescent prospermatogonia (Öllinger et al., 2008). *Tex19.1* may play a role in post-transcriptional silencing of retrotransposons during spermatogenesis, although further work is needed to clarify the role of *Tex19.1* in retrotransposon silencing in the germline.

1.9 Thesis outline

The general aims for the research project presented here were threefold:

1) To obtain a better idea about the extent of retrotransposon de-repression in *Tex19.1*$^{-/-}$ male germ cells and the mechanism by which *Tex19.1* might promote retrotransposon silencing in the male germline (chapter 3).

2) To analyse other aspects of the *Tex19.1*$^{-/-}$ knockout phenotypes in more detail. In chapter 4 I analyse the function of *Tex19.1* during embryonic development.

3) To characterise the phenotype of *Tex19.1*$^{-/-}$ females. This was the main focus of my PhD project and is presented in chapter 5.

Chapter 2: Materials and Methods

Most chemicals used throughout this project were of analytical grade and supplied by the manufacturer stated. Cell culture chemicals were cell culture grade and chemicals used for oocyte culture were of embryo grade.

2.1 Microbiological techniques

2.1.1 Growth of bacterial strains

For the preparation of plasmid DNA, competent DH5-α *Escherichia coli* (*E. coli*) (Invitrogen) transformed with the desired plasmid, were grown on either L-agar plates (10g/L NaCl, 10g/L Bacto-tryptone, 5g/L Yeast extract, 15g/L Difco Agar), inverted and incubated at 37°C overnight or in liquid culture (shaking at 200rpm) in Luria-Bertani (LB) medium (10g/L NaCl, 10g/L Bacto-tryptone, 5g/L Yeast extract). Both contained the appropriate antibiotic (Ampicillin at 50µg/L or Kanamycin at 50µg/L) to select for transformants carrying the desired plasmid. Growth of bacterial strains was carried out as described (Sambrook and Russel, 2001).

2.1.2 Bacterial transformations

For bacterial transformations 50µL aliquots of chemically competent DH5-α *E. coli* (library or subcloning efficiency) or One Shot TOP10 competent cells (Invitrogen) were incubated on ice with 10-100ng plasmid DNA for 30min. Cells were then heat shocked at 42°C for 30 seconds and returned onto ice for two minutes. 250µL of SOC medium (Invitrogen) was added to each transformation and the cultures incubated at 37°C for 60 minutes with shaking (200rpm) to allow recovery and expression of resistance genes. 20-100µL of the transformation mixtures was plated onto L-agar plates containing the appropriate antibiotic to select for the desired plasmid (Ampicillin at 50µg/µL or Kanamycin at 50µg/µL). Plates were incubated overnight inverted at 37°C to allow

growth and colony formation. In summary bacterial transformations were carried out according to the instructions provided by the manufacturer (Invitrogen).

2.2 Preparation and manipulation of DNA

2.2.1 Plasmid DNA isolation

For plasmid DNA isolation single bacterial colonies were picked from agar plates and used to inoculate 5mL cultures of LB medium containing the appropriate antibiotic for selection. Those were then grown in a 37°C shaking incubator. For mini-preparation of plasmid DNA, cultures were grown for 12-16 hours, cells pelleted by centrifugation at 4000g for five minutes and DNA isolated using QIAGEN spin miniprep columns (QIAGEN). For maxi-preparation of plasmid DNA, 5mL cultures were grown for eight hours and used to inoculate 100mL LB medium containing the appropriate antibiotic in a 1L conical flask. Cultures were grown with shaking overnight at 37°C and plasmid DNA was isolated by using QIAGEN maxiprep colums. Both isolation procedures were carried out following the instructions of the manufacturer.

2.2.2 Phenol/Chloroform extraction of nucleic acids

In principle nucleic acids were purified by phenol/chloroform extraction as described by Sambrook and Russel (2001). In brief, an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol (50% buffered phenol, 48% chloroform (v/v). 2% 3-methyl-1-butanol (v/v) pH >7.8) (Sigma) was added to the DNA/RNA preparation and vortexed for one minute before centrifugation at 10,000g in a benchtop microcentrifuge for ten minutes. The top aqueous layer was removed to a fresh tube, and the process repeated if a substantial interface was visible between the two layers after the first centrifugation. If not, an equal volume of chloroform was added to the nucleic acid

preparation, and the centrifugation step repeated. The aqueous layer was then subjected to ethanol precipitation to isolate DNA/RNA.

2.2.3 Ethanol precipitation of nucleic acids

Similarly to section 2.2.2, ethanol precipitation of nucleic acids was carried out according to the protocols provided by Sambrook and Russel (2001). For ethanol precipitations of nucleic acids 0.1 volumes of 3M sodium acetate pH 5.2 and 2 volumes of 100% ice-cold ethanol were added to DNA/RNA samples, mixed and incubated at –20°C over night. Samples were microcentrifuged at 10,000g in a benchtop microcentrifuge at 4°C for 15 minutes. The supernatant was discarded and the pellets washed with 70% ethanol and the centrifugation repeated. The supernatants were removed and pellets allowed to air dry at room temperature before resuspension in an appropriate volume of dH_2O or Tris-EDTA (TE) (10 mM Tris-Cl, pH 7.5. 1 mM EDTA) buffer.

2.2.4 Agarose gel electrophoresis

1-3% horizontal agarose (HiPure Low EEO Agarose, Biogene UK) gels (w/v) were prepared with 1x Tris/Borate/EDTA (TBE) buffer to resolve DNA samples. Ethidum Bromide (EtBr) was added to molten agarose to a final concentration of 0.5µg/mL before pouring to allow visualisation of DNA under ultraviolet light. 6x loading buffer (15% Ficoll 400 (Amersham Biosciences), 0.25% Orange G) was added to all DNA samples before loading. 500ng of 1kb ladder DNA size marker (Invitrogen) was loaded into a well on each gel to enable sizing and quantification of DNA fragments. Stained DNA was photographed using the Biorad Universal Hood II System (Biorad) and a thermal printer (Mitsubishi). Agarose gel electrophoresis was previously described by (Sambrook and Russel, 2001).

2.2.5 Gel purification of DNA fragments

PCR products or linearised plasmid DNA were separated from unwanted DNA fragments by agarose gel electrophoresis and isolated by gel purification in order to obtain clean products for cloning or other downstream applications. The appropriate band was excised from the gel using a clean razor blade under illumination from ultraviolet light. QIAQuick Gel Extraction Kit columns were used to isolate DNA fragments following the instructions provided by the manufacturer.

2.2.6 Analysis of nucleic acid quality and quantity

Quantification of DNA was measured either by comparison of the unknown quantity of DNA with a known quantity on an agarose gel or by measuring its absorbance at 260nm (A260) using a spectrophotometer by following the instructions provided by the manufacturer (Nandrop ND-1000, Nanodrop Technologies Inc).

2.2.7 Polymerase chain reaction (PCR)

PCR primers were designed by selecting a sequence between 17 and 25 nucleotides in length. Primers for use in RT-PCR reactions were, when possible, designed to span introns to prevent amplification of contaminating genomic DNA. Primer pairs were designed using the Primer 3 program (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) to reduce the chances of self-complementarity and therefore excess primer-dimer formation in the PCR reactions. Primer sequences are given in table 2.1. PCRs contained 0.5 units (U) Platinum Taq or Taq polymerase (Invitrogen), 0.5mM dNTPs (Invitrogen) and 1mM each primer (Sigma). 10x PCR buffer and 50mM Magnesium Chloride ($MgCl_2$) were supplied with the respective polymerase and used at a concentration of 1x and 2.5mM respectively. PCRs were conducted in a DNA Engine Tetrad PCR machine (MJ Reseach) in 0.1mL or 0.5mL

tubes. Typical PCR conditions used were 95°C for 5 min followed by 25 cycles of 95°C for 30 sec, 55 ± 10 ºC for 30 sec, 72°C for 1min. Reactions were then incubated for 72°C for 10 min. PCR has been described by (Sambrook and Russel, 2001).

2.2.8 Bisulfite conversion and sequencing of DNA

500ng genomic DNA was bisulphite treated using EZ Methylation Gold (Zymo research) and eluted in a final volume of 10µL according to manufacturer's instructions. Bisulphite primers were designed to amplify regions of interest (Table 2.1) and PCR amplification was performed in two rounds as follows: The first reaction was performed in a total volume of 50µL and contained 1x PCR buffer (Invitrogen), 2mM MgCl2, dNTPs (0.5mM each), 150µM each primer, 1.5 U platinum Taq polymerase (Invitrogen) and 1µL bisulphite treated DNA. Reactions were incubated in a DNA Engine Tetrad PCR machine (MJ Reseach) at 95°C for 5 min followed by 25 cycles of 95°C for 30 sec, 48ºC for 30 sec, 72°C for 1min. Reactions were then incubated for 72°C for 10 min. The PCR product was purified using a QIAGEN PCR purification column according to the instructions provided by the manufacture. 5µL of purified PCR product were used for the second amplification step. The second reaction was performed as described above but in a total volume of 100µL. Reactions were incubated in a thermocycler at 95°C for 5 min followed by 25 cycles of 95°C for 30 sec, 52ºC for 30 sec, 72°C for 1min. Reactions were then incubated for 72°C for 10 min. A negative control using 1µL water as template was performed for each set of reactions. The products were resolved on a 1.5% agarose gel. PCR resulted in one 450bp (full length) and one 350bp product. The desired product (full length band) was isolated performing gel extraction as described previously. Purified PCR products were ligated into pGEM Teasy vector (Promega) according to manufacturers instructions at 4°C overnight and transformed into DH5-□α *E. coli*; Library Efficiency (Invitrogen). Transformants were selected by plating bacteria on agar supplemented with ampicillin and X-gal. Up to 50 white colonies per transformation were cultured in 1mL LB supplemented with ampicillin and plasmid DNA was isolated and sequenced using SP6 sequencing primer (performed by

technical services, MRC HGU). Each sequence was imported into BioEdit program and aligned to a reference sequence for the region using ClustalW. The methylation status of each CpG dinucleotides from each bacterial clone was assessed by C to T conversion using QUMA online program (Kumaki et al., 2008).

Gene	Purpose of primer	Primer Sequences
Tex19.1	To genotype *Tex19.1* wild-type allele	5'-CTTCAGGAGGTCTGATGCCCTCT-3' 5'-GAGTGTTGTGTGGTGGGTGTTATGG-3'
Tex19.1	To genotype *Tex19.1* knockout allele	5'-CACCGCCTGTGCTCTAGTAGCTT-3' 5'-CTTCAGGAGGTCTGATGCCCTCT-3'
MMERVK10C	Bisulfite sequencing second amplification	5'-AGGTTTATAAAAGTAGTATTAC-3' 5'-ATAACCAATTAAAACAATAACATA-3'
MMERVK10C	Bisulfite sequencing second amplification	5'-TAAAAGTAGTATTAGTTTTGGC-3' 5'-AAACAAACAACACAATCCCA-3'

Table 2.1: PCR primer sequences.

2.3 Preparation and manipulation of RNA

2.3.1 RNA isolation and purification

RNA was isolated from cultured cells, embryoid bodies or tissues using TRIzol according to the instructions provided by the manufacturer (Invitrogen). Cultured cells were washed in ice-cold PBS and lysed in an appropriate amount of TRIzol. Cells were homogenized by passing the solution several times through a pipette followed by vortexing. Homogenized samples were incubated at room temperature to allow the dissociation of nucleoprotein complexes, before addition of 0.2mL chloroform per 0.75mL of TRIzol used, mixed and centrifuged at full speed in a benchtop microcentrifuge. The aqueous phase was collected, mixed with 0.5 volumes of isopropyl alcohol, per 0.75mL of TRIzol used, and incubated at -20°C over night and then centrifuged to precipitate the RNA. Pellets were washed with 70% ethanol prepared with RNase-free water (MilliQ). Air-dried RNA pellets were resuspended in 20 µl RNase-free dH_2O. The RNA was then treated with DNAse I by using the DNA-*free* kit from

Ambion in order to digest any possible genomic DNA contamination. This was carried out as recommended by the manufacturer. RNA was stored at –70°C.

2.3.2 cDNA synthesis

DNAse I treated RNA was used to produce cDNA with First Strand cDNA Kit (Invitrogen) according to the manufacturer's instructions. Before addition to the cDNA reaction mixture, RNA was heated to 65°C for ten minutes then placed on ice to remove secondary structure. To ensure maximum coverage of transcripts, random primers were used to generate random-primed cDNA. To check genomic DNA contamination was not occurring, reactions were set up with and without SuperScript III Reverse Transcriptase. 20µL cDNA reactions were prepared containing either 1µg of the corresponding total RNA, 250ng of random primers, 1µL 10 mM dNTP mix, 4µL 5x first strand buffer, 1µL 0.1M DTT, 1µl RNAse Inhibitor, 1µL SuperScript III Reverse Transcriptase, dH_2O up to 20µL. The reactions were incubated in a DNA Engine Tetrad PCR machine (MJ Reseach) at 25°C for 5min, 50°C for 60min and 70°C for 15min. cDNA preparations were stored at –20°C.

2.3.3 RT-PCR

To detect presence/absence of transcripts, PCRs were carried out using cDNAs as template. 1µL cDNA, or dH_2O as no template control, or 1µL of cDNA reactions set up without SuperScript II (Invitrogen) as no RT controls, were used as template in standard 20µL PCR reactions (2µL 10X PCR buffer, 0.6µl 50mM MgCl2, 0.3µl 10mM forward primer, 0.3µL 10mM reverse primer, 0.4µL 10mM dNTP mixture, 1U Taq polymerase. Primer sequences are given in Table 2.2. All reactions were subjected to thermal cycling in a DNA Engine Tetrad PCR machine (MJ Reseach) using the following conditions: 94°C for 2min, 25-35 cycles of (94C for 30sec, 48-55°C for 30sec, 72°C for 1min), 72°C for 5 minutes. 20µL of RT-PCR reaction products were resolved and photographed

on 2% agarose in 1x TBE gels as detailed above. PCR has been described by (Sambrook and Russel, 2001).

2.3.4 Quantitative-PCR

In order to detect and quantify the presence/absence of RNA transcripts quantitative RT-PCR (qPCR) was carried out using cDNAs as template. For each reaction cDNAs were diluted 1 in 50. For each reaction 5µL of a primer mix containing the required forward and reverse primers at a concentration of 1.2 pmol/µL were used. Controls reactions were also set up with dH_2O as no template control and dilutions of cDNA reactions set up without SuperScript II (Invitrogen) were used as no RT controls. qPCR was performed using SYBR Green PCR System (Stratagene) and a CFX96 Real-Time PCR Detection System (Bio-Rad). The conditions used for PCR were the following: 94°C for 15min, 45 cycles of (94°C for 15sec, 55°C for 30sec. 72°C for 30sec), melt curve from 65°C to 95°C with 0.5°C increment for 5min. Primers were validated to work under the used conditions at 90-100% efficiency. Primer sequences are given in Table 2.2. Three technical replicates were performed for each biological sample and the relative changes in gene expression determined using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). The transcript of the housekeeping gene ß-actin was used to normalise cDNAs and the expression levels relative to wild type ES cells or littermate controls are shown. A two-tailed t-test was used to determine statistical significance of qRT-PCR gene expression changes.

2.3.5 Illumina Beadarray Gene Expression Profiling of *Tex19.1*$^{-/-}$ testes

RNA was isolated from 16 dpp testes using TRIzol (Invitrogen) as described above. RNA was treated with DNAseI (Roche) for 2h at 37°C to remove genomic DNA contamination. cRNA samples were prepared using Illumina TotalPrep RNA

Amplification Kit (Ambion). Hybridisation to Illumina WG-6 arrays according to the manufacturers' protocols was performed by Wellcome Trust Clinical Research Facility, Edinburgh.

Gene	Purpose of primer	Primer Sequences
Oct4	*RT-PCR,qPCR*	5'-GTGGATTCTCGAACCTGGCT-3' 5'-GTCTCCAGACTCCACCCAC-3'
Msx3	*RT-PCR*	5'-TTGCGATTGGTTTTGTGTTT-3' 5'-CTCGAAGCCTGGAGCCTG-3'
Gata4	*RT-PCR*	5'-CCATCTCGCCTCCAGAGT-3' 5'-CTGGAAGACACCCCAATCTC-3'
Nkx2.5	*RT-PCR*	5'-GCTTTGTCCAGCTCCACT-3' 5'-AAGTGCTCTCCTGCTTTCC-3'
Nanog	*RT-PCR*	5'-TCCTCGCCCTTCCTCTGAA-3' 5'-CAGGACTTGAGAGCTTTTGTTTGG-3'
β-actin	*RT-PCR,qPCR*	5'-GGCTGTATTCCCCTCCATCG-3' 5'-ACATGGCATTGTTACCAACTGG-3'
Tex19.1	*RT-PCR, qPCR*	5'-AAAATGGGCCACCCACATCTC-3' 5'-CCACTGGCCCTTGGACCAGAC-3'
Tex19.2	*qPCR*	5'-ATAAGGCCTTGTTGGTCCTC-3' 5'-GGGGACACATGGTCAAGAAG-3'
Sdmg1	*qPCR*	5'-AGTGGCTTAAAGGAGACCATCA-3' 5'- TCAGCATGCGTTTCTCTATGTT-3'
MMERVK10C gag	*qPCR*	5'-CCTTTCCATGGTGTGATAGC-3' 5'-TGAATTTGTAGCTCGCATGA-3'
MMERVK10C pol	*qPCR*	5'-CCTTTCCATGGTGTGATAGC-3' 5'-TGAATTTGTAGCTCGCATGA-3'
MMERVK10C env	*qPCR*	5'-GGTAAAGTCTCCGAGGGTCA-3' 5'-AACTGGTCGCAGGAGCTG-3'
LINE1 ORF2	*qPCR*	5'-GGAGGGACATTTCATTCTCATC-3' 5'-GCTGCTCTTGTATTTGGAGCATAGA-3'
SINE	*qPCR*	5'-TGGTGGTGCATGCCTTTAAT-3' 5'-CCTGGTGTCCTGGAACTCACT-3'
IAP	*qPCR*	5'-GCACCCTCAAAGCCTATCTTAT-3' 5'-TCCCTTGGTCAGTCTGGATTT-3'
MERVL2a	*qPCR*	5'-GCCAGAGAGGTGCGGCAGTGGGC-3' 5'-GGACCCGTGGATCCTGGCTGTGGGA-3'
ETnERV2	*qPCR*	5'-ACAAATTCAGTATGGGCATC-3' 5'-GGGTACTGTTAAGACCCACA-3'
IAPEY3	*qPCR*	5'-ACAGAGGAGGACAACTGCTC-3' 5'-AACCTTACACAGGCAAAAGC-3'
MuLV	*qPCR*	5'-GGCAGCCATACATACAGACC-3' 5'-TGGTCTGCATAGAAACAGCA-3'
MysERV	*qPCR*	5'-GGACCAGACACCAAATAAGA-3' 5'-ATGTACAGGAAAGTGAAGG-3'
ERVB	*qPCR*	5'-GCAAAATGCCTTACTTCA-3' 5'-ACACAATCAGGACTGGTAG-3'
EtERV	*qPCR*	5'-AACACTTGCTGGTTTGTTTT-3' 5'-GATGGGGTGAGTAAAATTGA-3'
EtnI	*qPCR*	5'-TGAGAAACGGCAAAGGATTTTTGGA-3' 5'- ATTACCCAGCTCCTCACTGCTGA-3'
MusD	*qPCR*	5'-GTGCTAACCCAACGCTGGTTC-3' 5'-CTCTGGCCTGAAACAACTCCTG-3'

Table 2.2: RT-and qRT-PCR primer sequences.

2.4 Protein isolation and analysis

2.4.1 Total cell protein extracts from mammalian cells

Cells grown in monolayers were rinsed with phosphate buffered saline (PBS) and lysed in 20µL of PBS and 20µL of 2x SDS LaemmLi protein loading buffer (125mM Tris (pH 6.5), 4% SDS (w/v), 10% 2 β-mercaptoethanol (v/v), 20% glycerol (v/v), 0.1% bromophenol blue (w/v), 100mM DTT) per cm^2 of culture area, and scraped from the surface of the flasks. The cell lysates were passed several times through a 21-guage syringe and then boiled for 5 minutes. Samples were stored at –20°C until required.

2.4.2 Total cell protein extracts from mammalian tissues

Tissues were homogenized in 100µL of 2x Laemmli SDS protein loading buffer (see above) per 100mg of tissues with help of an electric pestle. Samples were then boiled for 5 minutes and stored at –20°C until required.

2.4.3 Resolution of proteins by SDS-PAGE

20µg of protein of each sample was resolved on polyacrylamide gels ((10% acrylamide (v/v), 0.39M Tris-HCl (pH8.8), 0.1% SDS (w/v), 0.1% ammonium persulphate (w/v), 0.04% TEMED (N,N,N',N-tetramethylethylene diamine) (v/v) in dH_2O) and stacking gels (4% acrylamide (v/v), 0.13M Tris-HCl (pH6.8), 0.1% SDS (w/v), 0.1% ammonium persulphate (w/v), 1% TEMED (v/v) in dH_2O)) using 30% acrylamide (29:1 acrylamide:bis-acrylamide (v/v) (Severn biotech). Gels were run in electrophoresis tanks (Hoefer) in Tris-gylcine running buffer ((25mM Tris base, 250mM glycine (pH8.3), 0.1% SDS (w/v)) at 40V/cm for around 1.5h. Pre-stained protein standards (Invitrogen) were loaded alongside samples to aid sizing of proteins. Preparation of SDS-gels for protein separation has been described by Sambrook and Russel (2001). Alternatively

20µg of protein of each sample was resolved on pre-cast Bis-Tris gels according to the instructions provided by the manufacturer (Invitrogen).

2.4.4 Western blotting

Western blotting was performed using standard procedure. After SDS-PAGE the stacking gel was removed and protein samples were transferred onto PVDF membrane (GE Healthcare) by wet western blotting using a GENIE blotter. Gels were transferred onto membranes within a sandwich of two pieces of Whatman paper above and below the gel and membrane. Gel, membrane and filter paper were soaked in pre-chilled transfer buffer (0.48M Tris, 39M Glycine, pH 9.2, 4°C) and placed between two electrodes. After 1.5h of transfer the stacks were disassembled and the membrane washed briefly in water. Membranes were rinsed in Ponceau's stain (Sigma) to visualise proteins on the membrane and allow assessment of the quality of the transfer. Membranes were then washed in PBS and blocked in blocking solution (5% (w/v) non-fat skimmed milk powder in PBST (PBS, 0.1% Tween-20)) for 20 minutes. Primary antibodies were diluted in blocking solution, and incubated with membranes overnight at 4°C with constant rotation. Three 5min washes in PBST preceded the incubation of the membrane with the secondary antibody in blocking solution for one hour at room temperature with constant rotation. Membranes were washed a further three times in PBST and detected using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific). Signals were exposed onto X-ray film (Kodak). Antibodies and dilutions used are given in table 2.3. If re-blotting was required membranes were stripped with a mild stripping buffer (0.1M glycine, pH2.1) twice for 15 minutes, blocked in blocking solution and incubated with primary and secondary antibodies as described above. Protocols for Western blotting were obtained from Sambrook and Russel (2001).

Antibody	Source	Species and concentration used
Tex19.1	Ian Adams' Group, Human Genetics Unit, UK	Polyclonal anti-rabbit 1:100
IAP	Bryan Cullen"s Group, Duke University, USA	Polyclonal anti-rabbit 1:1000
IAPE	Marie Dewannieux, Institute Gustave Roussy, France	Polyclonal anti-rabbit 1:500
LINE1 ORF1	Sandy Martin"s Group, University of Colorado, USA	Polyclonal anti-rabbit 1:2000
Pabp1	Nikki Gray"s Group, Human Reproductive Science Unit, UK	Polyclonal anti-rabbit 1:10000
β-actin	Abcam	Monoclonal anti-mouse 1:5000
Gapdh	Abcam	Polyclonal anti-rabbit 1:1000
Ubr2	Abacm	Monoclonal anti-mouse 1:1000
GFP	Roche	Monoclonal anti-mouse 1:2000
Oct4	BD Transduction Laboratories	Monoclonal anti-mouse 1:2000

Table 2.3: Antibodies used for western blot analysis.

2.5 Immunostaining

2.5.1 Immunohistochemistry (IHC)

IHC was basically performed as described (Öllinger et al., 2008). In brief, embryos were recovered from CD1 wild type matings, fixed at room temperature and fixed for 3 hours in 4% paraformaldehyde (PFA) in PBS. Afterwards the tissues were washed in PBS and

then dehydrated through ethanol, followed by xylene. The tissues were embedded in paraffin wax at 58°C. 6µm-thick sections were dewaxed in xylene and rehydrated. For antigen retrieval the slides were boiled for 15 minutes in 0.01 M sodium citrate, pH 6.0 in a microwave and then allowed to cool down to room temperature. The sections were blocked in PBS-BSA-azide containing 10% goat serum and 0.1% tween for 1 hour at room temperature. The slides were then incubated with 50µL rabbit anti-Tex19.1 (Ian Adams Group) primary antibody diluted in the blocking solution at 1:100 or 1:300. Anti-rabbit IgG (Sigma) served as negative control. Slides were washed in PBST. Bound antibody was detected using the DAKOvision ABC diaminobenzidine (DAB) kit as described by the manufacturer (DakoCytomation). Slides were counterstained haematoxylin as described by Puchtler et al., 1986.

2.5.2 Immunoflouresence (IF)

IF was basically performed as described (Öllinger et al., 2008). In brief, testes were collected from 16dpp *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$, 3 each. Tissues were fixed over night at 4°C in 4% PFA. Afterwards tissues were processed as described under 2.5.1. After antigen retrieval slides were blocked in PBS-BSA-azide containing 10% goat serum and 0.1% tween for 1 hour at room temperature. The slides were then incubated with 50µL rabbit anti-LINE1 ORF1 primary antibody alone or in combination with mouse anti-SYCP3 (Abcam) diluted in the blocking solution at 1:300 overnight at 4°C. Anti-rabbit IgG (Sigma) served as negative control. Slides were washed in PBST (PBS plus 0.01% Tween) incubated with green anti-mouse and red anti-rabbit Alexafluor secondary antibodies (diluted 1/500 in blocking solution, Invitrogen) for 1hr at room temperature, washed in PBST and then mounted with mounting media plus DAPI (250ng/mL).

2.6 Sucrose gradient analysis

To separate protein complexes cell lysates were centrifuged through a sucrose density gradient which was then fractionated, and the RNA or proteins in each fraction subjected to qPCR or western blotting.

2.6.1 Polysome gradient preparation

Polysome gradients were prepared as described by Gillian-Daniel et al. (1998). In brief, 18dpp testes were homogenized in 200µL of lysis buffer (20mM Hepes, 150mM KCl, 5mM DTT, 5mM $MgCl_2$, 100U/mL RNAsein, EDTA free mini complete protease inhibitor (Roche), 10nM calyculin A, ± 150µg/mL cycloheximide, ±20mM EDTA, ± 300µg/mL RNAse A) with help of an electric pestle. ES cells from 2 confluent 75cm^2 dishes were homogenized in 300µL of lysis buffer by pipetting up and down several times. After homogenization 0.5% of NP40 was added and the samples were then incubated on ice for 10 minutes. After lysing the cells nuclei and membranous debris were removed by centrifugation (12,000g, 5 minutes at 4°C). The supernatant was layered onto a 11mL linear sucrose gradient (10-50% sucrose w/v). For preparation of the gradient sucrose was dissolved in gradient buffer (20mM Hepes, 250mM KCl, 10mM $MgCl_2$, 1 µg/µL Heparin, 5mM DTT). Gradients were prepared either by freezing gradient method (layering of 1.67mL 50%, 42%, 34%, 26%, 18%, 10% on top of each other by freezing each individual concentration before applying the next one followed by defrosting over night at 4°C which facilitates formation of a linear gradient) or by using a gradient master (BioComp). Using the gradient master gradients were poured by mixing equal volumes of 10% and 50% sucrose solutions. Gradients were centrifuged in a SW41Ti rotor (Beckman) for 120 minutes at 38,000rpm at 4°C. After centrifugation 1mL fractions were collected and used for RNA or protein extraction. Fractions were collected with either a BiComp gradient station or a Pharmacia superfrac fraction collector. Absorbance of RNA at 254nm was recorded by using an econo UV monitor

(Bio-Rad) or inline UV monitor (Pharmacia) to assign location of polysomal, 80s, 40s and mRNP peaks to the fractions collected.

2.6.2 Size gradient preparation

Size gradient preparation was basically performed as described (Eskeland et al., 2004). To analyse at which molecular weight Tex19.1 migrates within a sucrose gradient, in order to examine if it might function in a complex, cytoplasmic ES cell cytoplasmic extract was loaded on 10-30% (w/v) sucrose gradients. Sucrose gradients were prepared in BC100 buffer (25 mM HEPES/KOH (pH7.3), 100 mM NaCl, 1 mM MgCl2, 0.5 mM EGTA, 0.1 mM EDTA, 10% glycerol (v/v), 1 mM DTT, and 0.2 mM PMSF) as described by Eskeland et al., 2004. ES cell extract was kindly provided by Ragnhild Eskeland. The gradient was prepared using a gradient master 105/106 (BioComp) set at 2.40 min/81.5 deg/15 rpm. Centrifugation was performed using a SW41 rotor (Beckman) at 41 000 rpm for 28 h at 4°C. 0.5 mL fractions were collected using the BioComp gradient station and analyzed by 10% SDS-PAGE.

2.6.3 RNA isolation from sucrose gradients

In order to isolate RNA from sucrose gradients, fractions were digested with 20µg/µL proteinase K in presence of 1% SDS and 10mM EDTA for 30 minutes at 37°C. RNAs were recovered by extractions with a 3 times volume of Trizol LS reagent according to the instructions provided by the manufacturer (Invitrogen). Isolated RNA was stored at –70°C until required. cDNA was synthesis was performed as described above.

2.6.4 Extraction of protein from sucrose gradient fractions

Proteins were isolated from sucrose gradient fractions by methanol precipitation. 600µL of methanol were added to 150µL of sucrose, followed by vortexing. Then 150µL were

added, samples were vortexed again followed by 450µl and a third vortex step. Samples were centrifuged immediately for 5min in a benchtop microcentrifuge. The upper aqueous layer was removed and discarded. Proteins were precipitated by addition of 650µL of methanol and centrifugation. Supernatant was removed, the protein pellet air-dried and then resuspended in Laemmli protein loading buffer and boiled for 5min. Isolated protein samples were stored at –20°C until required.

2.7 Mammalian cell culture

2.7.1 Freezing and thawing cells stored in liquid nitrogen

All cell suspensions were frozen in 1mL aliquots of 10% dimethyl sulfoxide (DMSO) (v/v), 20% fetal calf serum (FCS), and standard media (DMEM or GMEM) in screw-top cyrotubes (Nunc). To retrieve cells from liquid nitrogen, cells were thawed in a 37°C water bath as quickly as possible and then spun down in culture medium to remove the DMSO, before seeding into 25cm^2 culture flasks. Flasks for ES cell culture were pre-coated with sterile 0.1% gelatine in PBS for 15 minutes which was aspirated before addition of media and cells. Cell culture techniques have been described by Hogan et al., (1994).

2.7.2 Routine cell culture and harvesting

Feeder independent ES cells were cultured in E14 media (GMEM, 10% FCS, 1% non essential amino acids (NEAA), 1% sodium pyruvate, 1% penicillin/streptomycin (P/S), 1% L-glutamine (L-glut), 0.001% ß-mercaptoethanol (ß-ME), 1mL LIF conditioned media per 500mL GMEM, in a gelatinised culture flasks at 37°C until required confluence. Feeder dependent ES cells were cultured on fibroblast feeder layer. ES cells wild type, heterozygote and deficient for *Tex19.1* were cultured at 37°C in ES cell media (DMEM, 15% FCS, 1% P/S, 1% L-glut, 1% NEAA, 1% sodium pyruvate, 1µL LIF, 5µL

ß-ME). Primary embryonic fibroblasts or 3T3s were cultured in STO media (DMEM, 10% FCS, 10% P/S, 10% L-glut). To pass either of the cell lines above cells were washed once with PBS and then trypsinised for 2 minutes at 37°C. The reaction was stopped by adding 10x the amount of media containing FCS to the cells. Cells were spun at 1000rpm for 4 minutes, resuspended in fresh media, counted and either reseeded at a concentration of $2x10^5$/mL or frozen. To freeze the cells $8x10^6$cells/mL were resuspended in 500mL of media to which then 500mL of freezing media were added (DMSO, FCS, culture media). Cells were placed in polystyrene boxes at -70°C and after they were frozen stored at -150°C in liquid nitrogen. ES cell culture techniques have been described by Hogan et al., (1994).

2.7.3 Development of feeder cells for ES cell culture

Development of the feeder cells was performed as follows: Primary embryonic fibroblast cells were cultured until 80-90% of confluence in STO media (DMEM, 10% FCS, 10% P/S, 10% L-glut and then incubated with mitomycin C (MITC) (10µg/mL of MITC final concentration) at 37°C for 2.5 hours. After this time period cells were washed 3x with PBS, trypsinised and seeded at a density of $2x10^5$ cells/mL into gelatinised flasks. Cells were allowed to settle for at least 4 hours before used as feeders for ES cell culture. Generation of feeder cells for ES cell culture has been described by Hogan et al., (1994).

2.7.4 Cell counting

For cell counting, a drop of the resuspended cell suspension was placed on a haemocytometer with a weighted coverslip on top to generate an airtight seal. The total volume defined by the counting grid was $1x10^{-4}$ mL, and cell concentrations per mL were obtained by multiplying the total number of cells in the grid by $1x10^4$.

2.7.5 Embryoid body formation

ES cells were differentiated into embryoid bodies (EBs) by using hanging drop culture method. Ninety-six small hanging drops of ES cell media lacking LIF containing 600 cells per 20μL drop were cultured for two days suspended from the lids of bacterial petri dishes. The EBs that formed were transferred to bacterial culture dishes of 10 cm diameter and cultured in suspension in 10mL of ES media lacking LIF for a further five days. A half media change was performed every second day. At day 7 the EBs were transferred to a 10cm diameter tissue culture dish, onto which they attach, and form outgrowths in which extensive differentiation occurs. At day 10 beating structures were typically seen in the differentiated cultures indicating that cardiomyocytes had differentiated (Boheler et al., 2002). Samples of the embryoid bodies were taken at 0, 2, 4, 7 and 10 days of culture, RNA extracted and RT-PCR performed as described above to proof down regulation of *Tex19.1* upon differentiation

2.7.6 Generation of primary embryonic fibroblasts

Mouse embryonic fibroblasts were collected from embryos 12.5dpc of age, either from embryos obtained *Tex19.1*$^{+/-}$ matings in order to generate *Tex19.1*$^{-/-}$ and control fibroblasts or CD1 matings in order to generate feeder cells for ES cell culture. Placental and maternal tissues as well as head and all innards and blood vessels were removed from freshly harvested embryos. Carcasses were washed 3x in PBS containing 10% P/S. Three embryos per 75cm^2 culture dish were minced to make a slurry which was trypsinised in 1mL of trypsin for 10min. Trypsin activity was quenched by adding 1mL STO media and samples were pipetted up and down to break up tissues. Cells are allowed to grow up to confluency and splitted twice before being used in experiments, as feeders or frozen down as stocks. Derivation of feeder cells for ES cell culture has been described by Hogan et al., (1994).

2.7.7 Drug treatment of fibroblasts

Fibroblasts, either 3T3s or PEFs were seeded into T25 flasks at a concentration of $0.5x10^5$ cells per mL. The day after seeding, cells were treated with 1μM 5-aza-2'-deoxycytidine (Sigma), 1M of TSA (Sigma) or a combination of TSA and 5-aza-2'-deoxycytidine. The drug treatment was repeated every day for 3 days before the cells were harvested for RNA or protein expression analysis.

2.7.8 Transfection of cells

Cells were grown to 80% confluency and transfected with plasmid DNA using Lipofectamine (Invitrogen) diluted in OPTIMEN (Invitrogen) by following the instructions provided by the manufacturer. Slight amendments were made for transfection of ES cells which were transfected in ES media without serum to enhance transfection efficiency. Plasmids used for examination of N-end rule activity have been described by Dantuma et al., 2000. *Tex19.1* expressing plasmid was previously generated in the lab by Chao-Chun Hun by cloning the *Tex19.1* open reading frame into a pEXPR-IBA105 vector (IBA BioTagNology).

2.7.9 Luciferase assay

Following transient transfection, as described under 2.7.9, activities of firefly (Fluc) and Renilla (Rluc) luciferase were determined from ES cell lysates. Luciferase activity was measured using a Dual-Luciferase Reporter Assay System according to the instructions provided by the manufacturer (Promega), 40μL of cell lysates per assay and a Lumat LB 9507 luminometer (Berthold Technologies). Relative luciferase activity was obtained by dividing the firefly luciferase activity by the Renilla luciferase activity, then normalizing to the ratio calculated for the negative control, pRF with no insert, whose value was set

to 1. Plasmids were a kind gift of Sandy Martin from the University of Colorado. Generation and functional analysis of plasmids has been described by Li et al. (2006).

2.8 *In vitro* assays for protein/RNA interactions

2.8.1 mRNA capture assay

16dpp testes were homogenized in 100µl of lyses buffer per 100mg of tissues (20mM Hepes, 150mM KCl, 5mM DTT, 5mM $MgCl_2$, 100U/mL RNAsein, EDTA free mini complete protease inhibitor (Roche)) with help of an electric pestle. After lysis membranous debris was removed by centrifugation (12,000g, 5 minutes at 4°C). Those samples were then subjected to RNP capture with oligo(dT)-cellulose. Briefly, 10 mg of oligo(dT)-cellulose (Ambion) were blocked in 1 mL of binding buffer (20mM Hepes, 150mM KCl, 5mM DTT, 5mM $MgCl_2$) containing 5% BSA for 1 h at 4°C. Half of each sample was mixed with 0.5 mL of binding buffer and 25µL of blocked matrix. The other half was retained as input controls. For competition assays, 200 µg of $(A)_{25}$ (Sigma Oligos) was added 30 min before the addition of extracts. After being incubated for 1 hour min at 4°C on a rotating wheel, beads were spun down and washed three times with binding buffer; bound proteins were eluted by boiling in 40µL of 2× LaemmLi buffer. Ten percent of input and all of bound protein were western blotted for Tex19.1 and Pabp1 (see table 2.3).

2.9 Mouse husbandry

2.9.1 Animal care

The mice analysed in this study carry a targeted deletions of the *Tex19.1* gene in which the entire Tex19.1 open reading frame has been replaced with a neomycin cassette (Öllinger et al, 2008). Mice analysed here have been backcrossed 3x and maintained in a

C57BL/6 inbred background. C57BL/6 mice were obtained from Charles River. All animal work was conducted under a UK Home Office project licence with approval from an institutional ethics committee. Animal breeding and female copulation plug checking was conducted twice a day by standard methods. Females found plugged were scored as positive for insemination and marked as 0.5dpc. For superovulation females were stimulated with 5 IU pregnant mares' serum (PMS) and 5 IU human chorionic gonadotrophin (HCG) by intraperitoneal injection, 48 h apart, in order to induce superovulation. Mice were sacrificed by cervical dislocation.

2.9.2 Genotyping PCR

DNA was isolated from ear clips (taken from mice 12 dpp or older), embryo tail tips or whole embryos (up to 10.5dpc) by using DNAreleasy according to the instructions provided by the manufacturer (Anachem Ltd). Genotyping PCR was performed as described under 2.2.7. Sequences of genotyping primers are given in table 2.1. The wild type allele gives a band of 205bp and the knockout allele a band of 310bp.

2.9.3 Breeding analysis

Breeding data were obtained by genotyping pups born from *Tex19.1*$^{+/-}$ crosses. Analysis was done on pups born as first litters, pups born as subsequent litters from crosses where the male had been separated from the female after a plug was found and on pups born as subsequent litters from matings were the mating pair was constantly housed together in the same cage.

2.9.4 Chemically induced delay of embryo implantation and analysis of developmental potential

Embryonic diapause was induced in pregnant female mice at 2.5dpc by subcutaneous administration of tamoxifen (0.5µg per g of mouse in 100µL sesame oil) and intraperitoneal administration of depo-provera (0.05mg per g of mouse in 100µL PBS) as described by Hunter and Evans (1999). Diapause commences 2 days later, just before the normal time for implantation (4.5 dpc.). The ability of delayed blastocysts to resume development was investigated by flushing embryos from the uterus after two days of diapause and transferring them to pseudopregnant recipients at 2.5 dpc (this step was performed by Emma Murdoch). Pups born were analysed by genotyping as described above.

2.9.5 Lactation induced delay of embryo implantation and analysis of developmental potential

Tex19.1$^{+/-}$ mating pairs were set up and in order to ensure that the female conceived another litter during postpartum estrus only, the male separated from the female one day after the day the female had given birth. Lactation induced signaling resulted in embryonic diapause and its effects on development embryos was analysed by isolating the litters at the designated stages.

2.9.6 Female fertility data

Fertility data were obtained by mating *Tex19.1*$^{-/-}$ homozygote and *Tex19.1*$^{+/-}$ heterozygote females with C57BL/6 wild type males. On the day a plug was found mating pairs were separated and the females were allowed to carry the pregnancy to term. Fertility of females was assessed by the number of pups born from those matings.

2.10 Chromosome spreads

2.10.1 Chromosome metaphase spreads and counts of ES cells

Metaphase spreads were performed as described previously (Boyle et al., 2001). In order to enrich cell populations for cells with metaphase I chromosomes ES cells were cultured in presence of 0.1ug/mL colcemid (KaryoMAX) for 2 hours. Afterwards cells were harvested as described under 2.6.2 and washed 1x with PBS. Cells were incubated for 10min in presence of 0.25% KCl and 0.25% sodium citrate. Cells were then centrifuged at 900rpm for 5 minutes. Afterwards hypotonic solution was removed. Slowly 5mL of fixative (3:1 methanol/acetic acid; prepared fresh) was added drop wise to the cells. Fixative was removed by centrifugation and fresh fixative added. This procedure was repeated 3 times. Cells were dropped onto superfrosted glass slides from 10-20cm high. Slides were allowed to air dry. Chromosomes were then visualised by DAPI staining and analysed by microscopy. Remaining cells were stored at -20°C in fixative for extended times.

2.10.2 Pachytene spreads

Preparation and analysis of pachytene spreads from oocytes were performed as described previously (Bolcun-Filas et al., 2009). Ovaries were dissected from embryos and put into RPMI media (Invitrogen). The ovary was placed on a slide with one drop of RPMI and pricked with scalpels to release the oocytes. Sucrose (1.125g/ 25mL of H_2O) was added and oocytes were incubated for 1h, then treated with 0.05% Triton-X-100 for 10 mins and fixed in 2% PFA for 1h. Slides were washed in distilled H_2O and then blocked in PBS + 5% goat serum with 0.15% BSA and 0.1% Tween-20 for 1h at room temperature. Slides were incubated with the primary antibodies, mouse anti-Sycp3 (Abcam) and rabbit anti-Sycp1 (Abcam), diluted in blocking solution 1 in 200 and 1 in 250, respectively, overnight at 4°C. Slides were washed in PBST (PBS plus 0.01%

Tween) incubated with Alexa488 anti-mouse and Alexa594 anti-rabbit Alexa Fluor secondary antibodies (diluted 1/500 in blocking solution, Invitrogen) for 1hr at room temperature, washed in PBST and then mounted with mounting media plus DAPI (250ng/mL).

2.10.3 Metaphase I spreads

Females that were not in estrus were given an intraperitoneal injection of 5IU PMS between 3 and 3.30 pm. Ovaries were collected in KSOM media 42 hours later. Ovaries were then placed in 100µl prewarmed equilibrated media (alpha-MEM without nucleosides, containing 0.3mg/mL BSA, penicillin, streptomycin) and pierced repeatedly with a 25G needle to release oocytes. Oocytes were then transferred into a fresh drop of media, pipetted up and down to remove cumulus cells and washed 3x in fresh drops of media. GV-stage oocytes were selected and cultured at 37°C for 2 h. Oocytes that had undergone GV breakdown were selected and cultured for further 3 hours at 37°C. Chromosome spreads were performed as described by Yuan et al. (2002). After 5h of culture chromosome spreads were performed by incubating the oocytes in 1% sodium citrate for 15-20 minutes. Single oocytes were transferred to a superforst glass slide and fixed with 3:1 methanol:acetic acid fix. Slides were allowed to air dry and chromosomes stained with 0.05µg/mL DAPI in PBS for 10 minutes at room temperature. Slides were then mounted with Vectashield hard set mounting media, coversliped and analysed.

2.10.4 Metaphase II spreads

Metaphase II oocytes were harvested from super ovulated females 18 hours post HCG administration by dissecting out the oviduct and releasing zygotes from the ampulla into 37°C warm KSOM media. Oocytes were treated with 0.5mg/mL hyaluronidase for five minutes to remove cumulus cells and then washed 3x in FHM. Chromosomes were spread as described above and after fixation stained by DNA FISH using a major

satellites probe in order to facilitate chromosome counting. FISH was carried out as described previously (Boyle et al., 2001). For DNA FISH the slides were washed briefly in 2× SSC and incubated with 100µg/mL RNaseA for 1 h at 37°C. The slides were dehydrated using 70, 90, 100% ethanol and air dried. Slides were denatured in 70% formamide/2× SSC at 70°C for 75s, transferred to ice-cold 70% ethanol, and then into 90 and 100% ethanol and air dried. Fifteen microlitres of 100ng biotin-labelled major satellite paint were denatured at 70°C, re-annealed at 37°C for 15 min and hybridized on the slide under a sealed coverslip overnight at 37°C.). Probes were kindly provided by Shelagh Boyle. The slides were washed 4× 3 min in 2× SSC at 45°C, 4× 3 min in 0.1× SSC at 60°C and then 4× SSC 0.1% Tween-20 at room temperature. Biotinylated probes were detected using FITC-conjugated avidin, followed by biotinylated anti-avidin and a final layer of FITC-conjugated avidin (Vectashield). Slides were mounted with Vectashield hard set mounting media, containing DAPI, coversliped and analysed.

2.10.5 Anaphase II spreads

Metaphase II oocytes were harvested from super ovulated females 16 hours post HCG administration by dissecting out the oviduct and releasing zygotes from the ampulla into KSOM media. Oocytes were treated with hyaluronidase to remove cumulus cells and then washed 3 times in FHM media. Oocytes were then cultured in FHM media containing 5% P/S, 5mM $SrCl_2$ (Sigma) and 2mM EGTA (Sigma) for 2h as described by Kishigami and Wakayama (2007). Chromosome spreading and analysis was performed as described under 2.10.3.

2.10.6 Zygote chromosome spreads

Zygotes were harvested from naturally mated plugged females at 0.5dpc by dissecting out the oviduct and releasing zygotes from the ampulla into KSOM. Zygotes were treated with hyaluronidase as described under 2.10.4 to remove cumulus cells and then washed 3 times in FHM. Zygotes were then incubated in KSOM with 0.1µg/mL

colcemid overnight. Chromosome spreading and analysis was performed as described under 2.10.3.

2.11 Microscopy

2.11.1 Brightfield microscopy

Brightfield images were taken using a Coolsnap HQ CCD camera (Photometrics Ltd, Tucson, AZ) Zeiss Axioplan II fluorescence microscope with Plan-neofluar objectives (Carl Zeiss, Welwyn Garden City, UK). Colour additive filters (Andover Corporation, Salem, NH) installed in a motorised filter wheel (Ludl Electronic Products, Hawthorne, NY) were used sequentially to collect red, green and blue images, which were then superimposed to form a colour image. Image capture and analysis were performed using in-house scripts written for IPLab Spectrum (Scanalytics Corp, Fairfax, VA).

2.11.2 Fluorescence microscopy

Fluorescence images were taken using a Hamamatsu Orca AG CCD camera (Hamamatsu Photonics (UK) Ltd, Welwyn Garden City, UK) Zeiss Axioplan II fluorescence microscope with Plan-neofluar objectives, a 100W Hg source (Carl Zeiss, Welwyn Garden City, UK) and Chroma #83000 triple band pass filter set (Chroma Technology Corp., Rockingham, VT) with the excitation filters installed in a motorised filter wheel (Prior Scientific Instruments, Cambridge, UK). Image capture and analysis were performed using in-house scripts written for IPLab Spectrum (Scanalytics Corp, Fairfax, VA).

2.12 Statistical analysis

Chi-square test was used to test for statistical significant difference in distribution of categorical data when there were more than two categories of data. Two-tailed Fisher's exact test was used to analyse significance in cases were sample sizes were small ($n < 5$) and/or the number of categories was two. Mann-Whitney-U test, a non-parametric test, was used to analyse for significant difference between variables where a normal distribution could not be assumed. Unpaired two-tailed Student's *t*-test was used to analyse significance of differences in continuous data between groups. Statistical significance was considered as $p < 0.05$. Statistical analyses were performed using Excel or GraphPad software (GraphPad Software Inc, La Jolla, CA USA).

Chapter 3: Investigation of the role of the Tex19.1 protein in retrotransposon repression

3.1 Introduction

The first aim of this thesis was to gain a better understanding of the mechanism by which *Tex19.1* promotes retrotransposon silencing. As described in chapter 1, *Tex19.1*$^{-/-}$ meiotic spermatocytes show de-repression of a specific retrotransposon called *MMERVK10C*, the upregulation of which was identified by microarray (Öllinger et al., 2008). The rational for performing a microarray in this study was that *Tex19.1* might influence the expression of genes known to be required for meiotic chromosome synapsis. However, the microarray showed no indication of expression changes in genes that are known to be required for spermatogenesis, but instead revealed upregulation of five probes in the *Tex19.1*$^{-/-}$ tissue, all of which were not annotated with gene symbols (Öllinger et al., 2008). The corresponding probe sequences gave multiple BLAST hits scattered throughout the mouse genome, which suggested they might recognize repetitive elements. Aligning these probe sequences against the Repeatmasker database of repetitive elements revealed that these probes were 100% identical to the ERVK LTR retrotransposon *MMERVK10C*. This observation was validated by qRT-PCR, Northern blotting and in situ hybridization (Öllinger et al., 2008).

Öllinger et al. (2008) did not observe an upregulation of *LINE1*, *SINE* or *IAP* elements in *Tex19.1*$^{-/-}$ testes. Carrying on from this work, one of the aims of this chapter was to clearly establish if other retrotransposons are upregulated in the absence of *Tex19.1* in testes, or if the de-repression is indeed specific to *MMERVK10C*. Furthermore, we wanted to understand how loss of *Tex19.1* could affect *MMERVK10C* expression. In contrast to the frequently studied and well characterised *IAP* or *LINE* retrotransposons little was known, at the beginning of the project, about how expression of *MMERVK10C* is generally controlled in somatic, germ or pluripotent cells. As described under 1.7.2 employ several different cell types DNA methylation and/or H3K9 methylation to control retrotransposon transcription. In order to obtain a better overview of the transcriptional regulation of *MMERVK10C*, and hence a better understanding of why

MMERVK10C might be up-regulated in *Tex19.1*$^{-/-}$ testes, we analysed the RNA expression levels of *MMERVK10C* in different DNA methylation and/or H3K9 methylation cell line mutants.

However, mis-expression of *MMERVK10C* RNA could not only be caused by failure of transcriptional repression but by failure to inhibit any stage of the retroviral life cycle. For example increased retroviral protein stability might result in increased retrotransposition events which would introduce more copies of the element into the genome which would in turn result in higher levels of RNA transcripts produced. It is also possible that an increase in translation might cause increased stability of *MMERVK10C* transcript resulting in the upregulation of *MMERVK10C* RNA observed in *Tex19.1*$^{-/-}$ mutants as decay and translation are two competing fates for cytoplasmic mRNAs (Shyu et al., 2008). Translation of an mRNA itself occurs in large protein structures which are called the polysomes (Ruan et al., 1997). Sucrose gradients can be used to gain an indication of the state of an mRNA molecule in the cytoplasm (Ruan et al., 1997). Actively translated mRNAs loaded with ribosomes are found in the heavier compartment of a sucrose gradient. As can be seen from figure 3.13 polysomes are readily separated according to the number of associated ribosomes (Esposito et al., 2010). The 40S and 60S peaks represent mRNAs bound by either the small or large ribosomal subunits, respectively (figure 3.1). Joining of the 40S and the 60S subunits forms the 80S ribosome, mRNAs attached with one ribosome are also referred to as monosomes, and translation is initiated when the assembled ribosome identifies and commits to a translational start codon (Jackson et al., 2010). mRNAs that are not translated form messenger ribo-nucleo-protein (mRNP) particles which can be distinguished on a sucrose gradient from the translated transcripts as they sediment slower than mRNAs associated with ribosomal subunits or whole ribosomes (figure 3.1) (Ruan et al., 1997). Translational regulators of mRNAs might be expected to be found in the polysome compartment of a sucrose gradient where they might stall translating ribosomes on mRNAs or induce premature termination of translation. Very potent

translational repressors might prevent access of the ribosomal subunits to mRNAs in the first place and therewith sequester them away from the translational machinery.

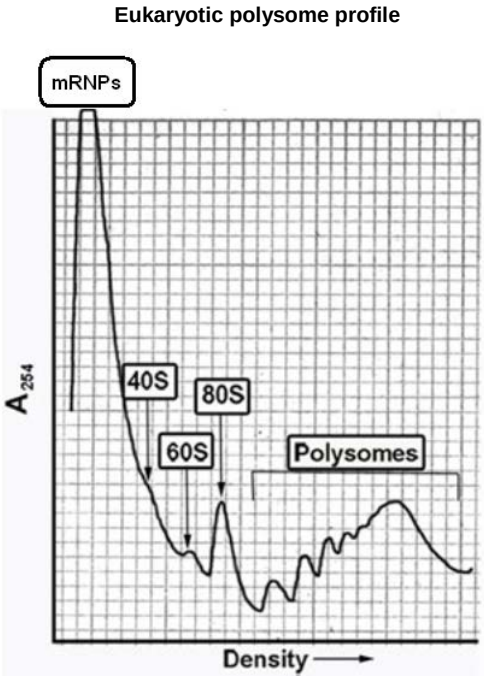


Figure 3.1 Representative trace of a eukaryotic polysome profile. Peaks containing mRNPs, small and large ribosomal subunits (40S and 60S, respectively and polysomes are indicated. Figure was taken and modified from Esposito et al. (2010).

Ian Adams utilised sucrose gradients to asses the state of *MMERVK10C* mRNA in the cytoplasm of *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ male germ cells. He found that *MMERVK10C* mRNA migrates in the heavier fractions of a sucrose polysome gradient, the compartment where mRNAs are more efficiently translated, in *Tex19.1*$^{-/-}$ testis compared to *Tex19.1*$^{+/+}$ controls (figure 3.2). This result raised the possibilities that Tex19.1 might either promote the degradation of actively translated *MMERVK10C* mRNA or be acting as a translational repressor of the *MMERVK10C* retrotransposon.

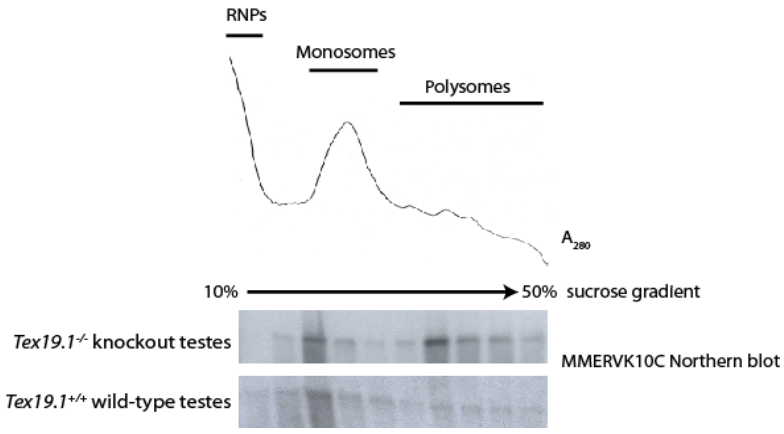


Figure 3.2 Northern blot for MMERVK10C transcript from 16dpp *Tex19.1*$^{-/-}$ and control testes sucrose gradient fractions (Ian Adams).

In this chapter I carry on from those initial observations by determining the localisation of Tex19.1 protein on a sucrose polysome gradient as well as further characterising the localisation of retrotransposon mRNA in *Tex19.1*$^{-/-}$ and control testes.

Another indication for a possible mechanism by which *Tex19.1* may repress retrotransposons came from a recently published paper by Jeremy Wang's group (Yang et al., 2010). The data they present suggest that Tex19.1 acts on the protein rather than on the mRNA level as they identified Ubr2 as a binding partner of Tex19.1 in wild type testes (Yang et al., 2010). This finding was of great interest as *Ubr2*$^{-/-}$ mice to some extent phenocopy *Tex19.1*$^{-/-}$ mice. *Ubr2*$^{-/-}$ mice exhibit, like *Tex19.1*$^{-/-}$ mice, defects in the progression through meiotic prophase in male germ cells and show a reduction in the numbers of *Ubr2*$^{-/-}$ females born (Kwon et al., 2003; An et al., 2010). Ubr2 itself is a ubiquitously expressed E3 ubiquitin ligase implicated in the N-end rule pathway of protein degradation. The N-end rule pathway states that the nature of the N-terminal amino acid of a protein, also called N-degron, defines the half-life of the protein (Tasaki and Kwon, 2007). Amino acids are classified as stabilizing or destabilizing residues depending on whether they increase or decrease the likelihood of a protein being degraded (figure 3.3). E3 ubiquitin ligases that function in the N-end rule pathway, like Ubr2, recognize destabilizing residues and label substrates for proteolysis by covalently

linking an internal Lys residue of the substrate with an ubiquitin moiety. Ubiquitin plays an essential role as a secondary signal for proteolysis in eukaryotes by targeting proteins for degradation by the 26S proteasome (figure 3.4) (Bedford et al., 2010). The selectivity of ubiquitylation according to the N-end rule pathway represents the crucial step in substrate selection and is mainly determined by the E3 enzymes (N-recogins) (Varshavsky, 1997). Based on the physical interaction between Tex19.1 and Ubr2, found by immuno-precipitation from testes extracts, as well as on the phenotypic similarities between *Ubr2*$^{-/-}$ and *Tex19.1*$^{-/-}$ mice, we reasoned that both proteins might function as a complex in the same pathway. In order to understand if Tex19.1 might affect retrotransposon protein stability in this chapter I examined if Tex19.1 might promote N-end rule mediated protein degradation.

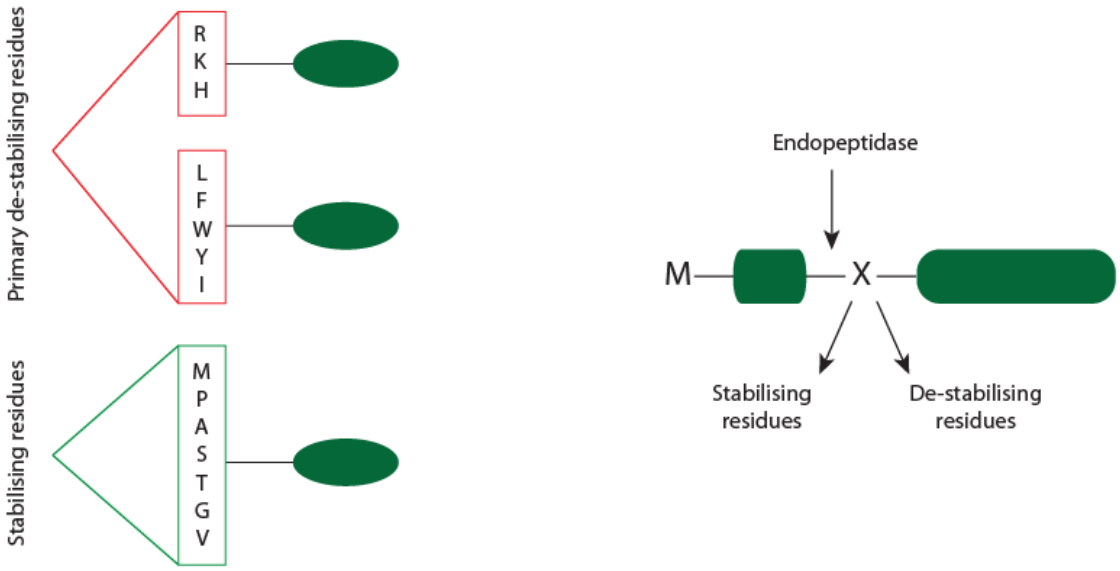


Figure 3.3 The N-end rule code and generation of N-degrons. In eukaryotes the basic residues Arg, Lys and His and the hydrophobic residues Leu, Phe, Trp, Tyr and Ile function independently as destabilizing residues. Met, Pro, Ala, Ser, Thr, Gly and Val are stabilizing residues. Endoproteolytic cleavage of proteins results in protein fragments with novel N-termini which can include stabilizing and destabilizing residues. Modified from Mogk et al. (2007).

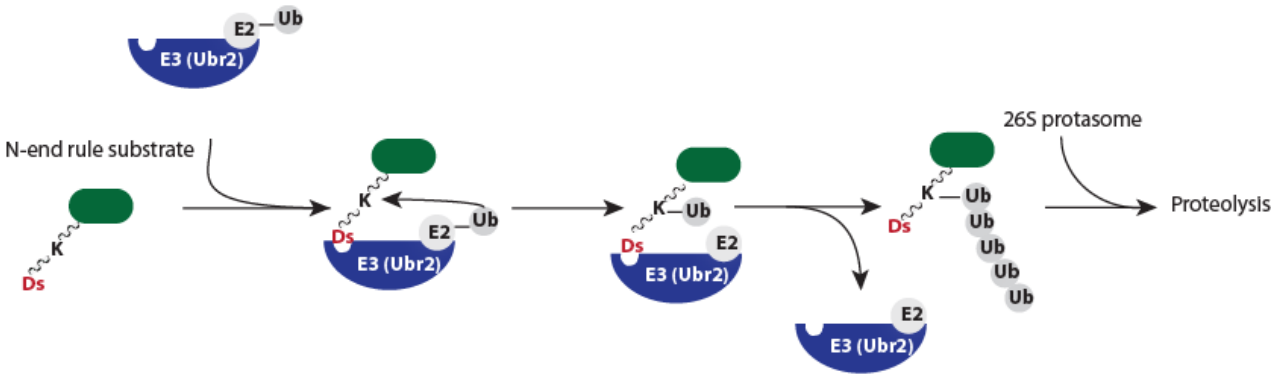


Figure 3.4 The N-end rule pathway of protein degradation. In eukaryotes destabilizing residues (Ds), also called N-degrons, of the protein substrate (green) are recognized by E3 ligases (blue) which contain specific binding sites for the destabilizing residue (Ds). The E3 ligase causes in combination with an E2 ubiquitin-conjugating enzyme (E2) the attachment of ubiquitin (Ub) to a lysine (K) on the target protein. This is followed by attachment of a second ubiquitin to the first, of a third to the second and so on. The multi-ubiquitylated protein substrate is recognized by the proteasome and targeted for proteolysis. Modified from Mogk et al. (2007).

3.2 Results

3.2.1 Retrotransposon mRNA expression in *Tex19.1*$^{-/-}$ testes

3.2.1.1 Analysis of retrotransposon RNA expression in *Tex19.1*$^{-/-}$ testes

The analysis of individual upregulated probes in the previously performed microarray in *Tex19.1*$^{-/-}$ testes lead to the identification of *MMERVK10C* as being upregulated in the absence of *Tex19.1* in mice on a mixed genetic background (129/Ola x CD1) (Öllinger et al., 2008). During this project we wanted to take this finding further and examine how widespread retrotransposon de-repression in *Tex19.1*$^{-/-}$ mutants might be. In order to do so I performed gene expression profiling of mutant and control mice, in which the *Tex19.1* mutation was backcrossed onto an inbred C57BL/6 genetic background, in order to minimise genetic variation between the animals, using the repeat-annotated Illumina Mouse WG-6 v2.0 Beadchip probes. Prepubertal mice at the age of 16dpp were used as the defects in progression through meiosis, which perturb the normal cellular composition of the testis, are first becoming apparent at this stage and are therefore unlikely to have a major influence on the results (Öllinger et al., 2008). In order to monitor repetitive element expression on a genome-wide scale Illumina Mouse WG-6 v2.0 Beadchip probes were repeat annotated and the obtained gene expression profiles analysed by Ian Adams as described (Reichmann et al., submitted). Ian Adams identified that 2.6% of probes on the array (494 out of 19,089) mapped to repeats and could be classified into 172 different repetitive elements (including LTR retrotransposons, *LINE*, *SINE* and DNA transposons) which in general, similar to other probes on the array, showed no significant changes in expression levels between *Tex19.1*$^{-/-}$ and control animals (Reichmann, et al., submitted). However, ten probes (0.05%) on the array were significantly upregulated by at least 2-fold in *Tex19.1*$^{-/-}$ testes and six of these probes belonged to the LTR retrotransposon class all of which were identified to map to *MMERVK10C*. In order to verify our results obtained from the repeat-annotated Illumina

Beadarray data I performed qRT-PCR on 16dpp *Tex19.1*$^{-/-}$ and control testes. I was able to confirm an approximately 2.2-fold upregulation of *MMERVK10C* transcript by qPCR in *Tex19.1*$^{-/-}$ testes compared to controls (figure 3.4A) (two-tailed t-test, $p<0.05$). I analysed the expression of *MMERVK10C* RNA in *Tex19.1*$^{-/-}$ testes in more detail by performing qRT-PCR for *MMERVK10C gag*, *pol* and *env* transcripts, which encode the three major proteins within the retroviral genome, by using primers specific for either sequence. I found that both *MMERVK10C env* and *pol* transcripts were significantly upregulated in *Tex19.1*$^{-/-}$ mice (two-tailed t-test, $p<0.05$) (figure 3.4B). In contrast, *MMERVK10C gag* expression appeared not to be different between mutant and control testes (two-tailed t-test, $p>0.1$) (figure 3.5B). This is consistent with the microarray data which suggested that *MMERVK10C env* probes were upregulated to a larger extent than the probe complementary to the *gag* region of *MMERVK10C* (Reichmann et al., submitted). The *pol* sequence is not covered by probes on the array. I also performed qPCR for a variety of other retrotransposons that were represented by probes on the microarray but did not appear to be upregulated in *Tex19.1*$^{-/-}$ testes. I was able to confirm that *LINE1* as well as one representative member of each the *ERV1* (*RLTR4*), *ERVK* (*IAP*) and *ERVL* (*MERVL2a*) LTR retrotransposon families do not change in expression in the absence of *Tex19.1* (figure 3.4.A). Similarly, *IAPEY3* and *MMERVK9E*, two retrotransposons closely related to *MMERVK10C*, show no change in expression in *Tex19.1* mutants (figure 3.4.A). Taken together, systematic annotation and analysis of repeat-annotated Illumina Mouse WG-6 v2.0 microarrays and subsequent qPCR confirm the *MMERVK10C* upregulation in *Tex19.1*$^{-/-}$ C57BL/6 testes, identified previously on a mixed (129/Ola x CD1) genetic background, and extend the range and variety of repetitive elements expression analysed in *Tex19.1*$^{-/-}$ mutants. Intriguingly, among the 172 expressed elements represented in this microarray dataset, *MMERVK10C* constitutes the only repetitive element whose expression changes by more than 2-fold in the absence of *Tex19.1*. Furthermore, the upregulation of *MMERVK10C* appears to be fairly specific as even closely related retrotransposons are not upregulated in *Tex19.1*$^{-/-}$ testes compared to controls. This raises the question why specifically *MMERVK10C* is de-repressed in *Tex19.1*$^{-/-}$ males and if inhibition is lost at the DNA, RNA or protein

level when *Tex19.1* is absent as mis-expression of *MMERVK10C* RNA could be caused by failure to suppress any stage of the retroviral life cycle.

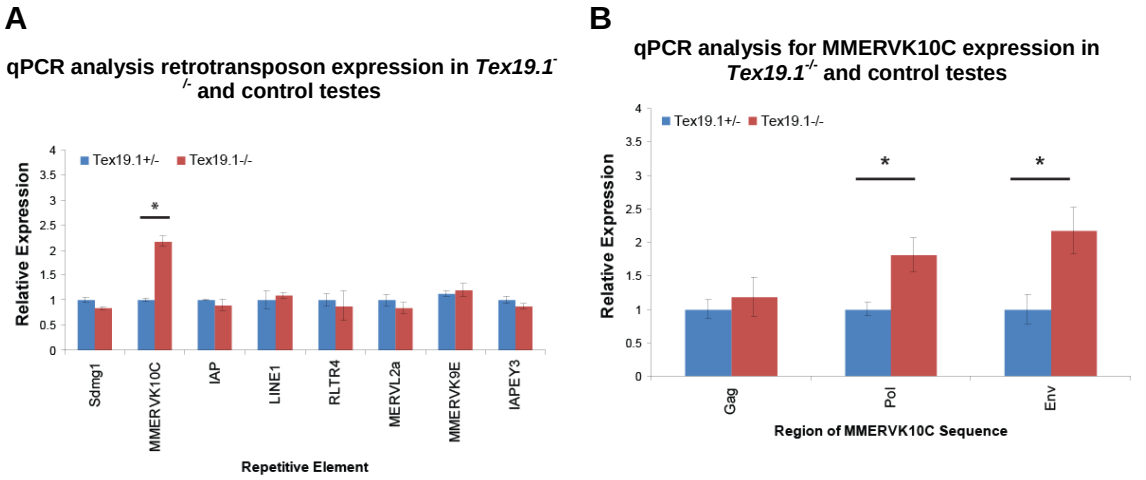


Figure 3.4 qRT-PCR showing relative expression of repetitive elements in *Tex19.1*$^{-/-}$ knockout and control 16dpp testes. (A and B) Expression levels for each repetitive element (mean ± standard error for three animals) were normalized to *β-actin* and expressed relative to littermate controls. *Sdmg1* is a single-copy control gene expressed in sertoli cells, which should not be affected in *Tex19.1*$^{-/-}$ mice, to verify normalization between animals. Asterisks indicate statistically significant differences (two-tailed t-test, $p<0.05$).

3.2.2 Transcriptional regulation of *MMERVK10C*

3.2.2.1 Transcriptional regulation of *MMERVK10C* by DNA methylation

As described in chapter 1, it is well established that DNA methylation plays a pivotal role in transcriptional silencing of certain retrotransposons. For example the well characterised *IAP* elements, which belong to the same LTR retrotransposon family (ERVK) as *MMERVK10C*, have been previously reported to be repressed by DNA methylation in somatic cells and germ cells (Jackson-Grusby et al., 2001; Bourc'his and Bestor, 2004). However, *MMERVK10C* is less well studied and at the beginning of the project no information was available concerning how expression of this relatively obscure element is regulated. Previous work in our laboratory could not detect any notable changes in DNA methylation status of *MMERVK10C* in *Tex19.1*$^{-/-}$ testes

(Öllinger et al., 2008). It can not be excluded though, given the presence of large numbers of *MMERVK10C* copies in the mouse genome that defects in DNA methylation status might have been missed in this analysis or that only specific copies might be affected. Therefore we took a more general approach in order to identify how *MMERVK10C* transcription can be silenced.

To understand if DNA methylation might contribute to *MMERVK10C* repression I analysed expression of this element in mouse embryonic fibroblasts (MEFs) which are hypomorphic for *Dnmt1* (I will refer to those as *Dnmt1*$^{-/-}$) and represent a model for genome wide hypomethylation (Lande-Diner et al., 2007). Null mutations in *Dnmt1* normally result in apoptosis in somatic cells - the cells used here were double mutants for *Dnmt1* and *p53* with the controls being only mutant for *p53* (Lande-Diner et al., 2007). For this experiment, *Tex19.1* and *IAP* serve as positive controls as RNA upregulation of those two genes in response to hypomethylation, has previously been reported (Walsh and Bestor, 1999; Gaudet et al., 2004, Hackett et al., submitted). *Tex19.1* mRNA expression is elevated by ~ 450-fold and *IAP* transcript by ~ 50-fold in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs. In contrast, *LINE1* (downregulated by ~1.6-fold), *MusD* (~1.4-fold change) and *MMERVK10C* (~1.3-fold change) respond less strongly to mutations in *Dnmt1*$^{-/-}$ in fibroblasts. Statistical significance of those changes could not be assessed as we only had one *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ and one control line available for those experiments. In order to circumvent this caveat I performed the same analysis on experimentally hypomethylated fibroblasts (figure 3.6). However, the data presented in figure 3.5A suggest that other and/or additional mechanisms may contribute to the transcriptional silencing of those elements in somatic cells (figure 3.5A). To test if DNA methylation levels were decreased in *Dnmt1*$^{-/-}$ MEFs I performed bisulfite sequencing of the *MMERVK10C* LTR/5'UTR region which revealed a ~ 50% reduction in DNA methylation in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ knockout compared to *p53*$^{-/-}$ control MEFs (figure 3.5B). It is worth noticing that the element is heavily methylated in the first place (figure 3.5B). The extent of de-methylation of *MMERVK10C* is lower compared to i.e. the *Tex19.1* promoter which shows around 80% reduction in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ knockout compared to

p53$^{-/-}$ control fibroblasts (J. Reddington, personal communication). The lower reduction in methylation levels at *MMERVK10C* loci compared to the *Tex19.1* locus could be explained by active targeting by *de novo* methyltransferases (Jones and Liang, 2009). Alternatively, the sensitivity to the hypomorphic allele of *Dnmt1* might differ between genomic loci. The primers used for bisulfite sequencing and qPCR analysis of *MMERVK10C* were designed to pick up a large proportion of the elements in the mouse genome. Given the large number of copies in the mouse genome it can not be excluded that a subset of *MMERVK10C* elements, which are not recognised by the qPCR primers used, might be upregulated in hypomethylated cells. However, the data I present suggest that the majority of *MMERVK10C* elements are not upregulated upon loss of methylation in fibroblasts.

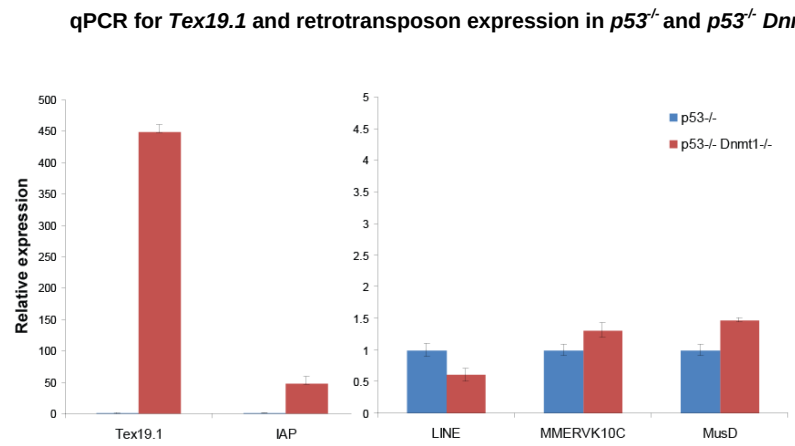
Given the result that *MMERVK10C* RNA appears not to be upregulated in hypomethylated fibroblasts I hypothised based on

- the observation that *MMERVK10C* is upregulated in a germ cell context when *Tex19.1* is absent and
- the fact that *Tex19.1* is mis-expressed in hypomethylated fibroblasts

that *Tex19.1* might be one of the factors contributing to *MMERVK10C* silencing in the event of global hypomethylation. To address this I experimentally hypomethylated *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/+}$ primary MEFs (pMEFs) and performed qPCR for expression of *MMERVK10C*. The reasoning behind this experimental design was that if *Tex19.1* should be involved in *MMERVK10C* silencing in hypomethylated somatic cells *Tex19.1*$^{-/-}$ pMEFs would be expected to show a larger upregulation of *MMERVK10C* transcript than *Tex19.1*$^{+/+}$ pMEFs when experimentally hypomethylated. I induced hypomethylation during culture by treating the cells with 5'-deoxy-2'-azacytidine (Aza). Aza is an analogue of cytosine and has DNA demethylating effects but also shows cytotoxicity at a certain dose (Patel et al., 2010). It becomes incorporated into DNA during S-phase and covalently links Dnmts to DNA. This creates genome-wide Dnmt protein–DNA cross-links which results in a reduction in the soluble Dnmt1 protein pool (Patel et al., 2010). Furthermore, Aza causes proteasomal degradation of non-chromatin

bound Dnmt1 through a mechanism which depends on DNA synthesis and targeting of DNA incorporated Aza residues by Dnmt1 itself (Patel et al., 2010). Both effects lead to replication-dependent global hypomethylation.

A



B

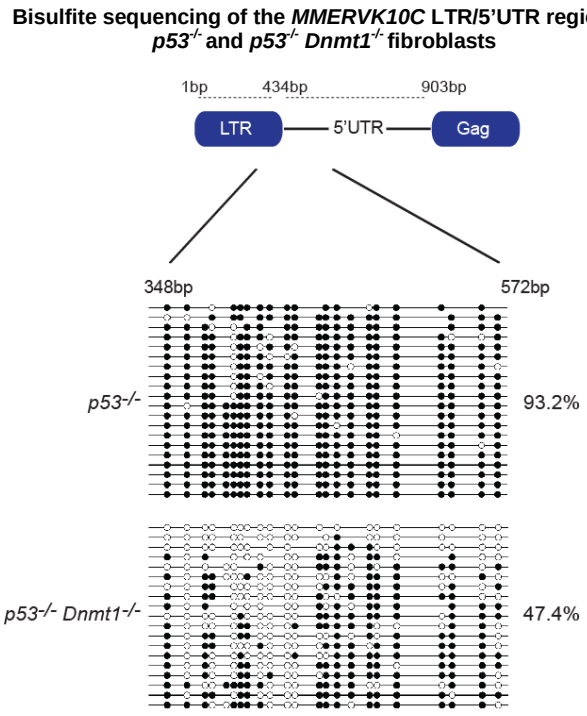


Figure 3.5 RNA expression of repetitive elements and *Tex19.1* and methylation status of *MMERVK10C* in hypomethylated fibroblasts. (A) Expression levels for each repetitive element and *Tex19.1* (mean ± standard error for 3 technical replicates) were normalized to *β-actin* and expressed relative to *p53*$^{-/-}$ controls. (B) Bisulfite sequencing showing the methylation

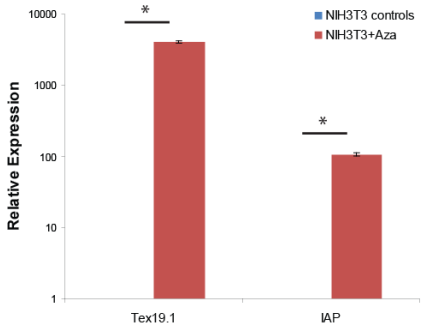
status of *MMERVK10C* LTR/5'UTR in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ and *p53*$^{-/-}$ control fibroblasts. Bp are shown relative to LTR start site (LTR = 1bp to 434bp; 5'UTR = 435bp to 903bp).

First I treated NIH3T3 mouse embryonic fibroblasts with Aza in order to strengthen the observations regarding gene expression made in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ fibroblasts and to analyse if *Tex19.1* RNA would get translated into protein in somatic cells. Figure 3.6A shows that treatment of NIH3T3 fibroblasts results in a ~ 4000-fold increase of *Tex19.1* and ~ 100-fold increase of *IAP* expression compared to controls (two-tailed t-test, $p<0.005$) (figure 3.6A). As these are genes are known to be silenced by DNA methylation the conditions used here seem sufficient to induce Aza dependent hypomethylation in NIH3T3 fibroblasts (Jackson-Grusby et al., 2001; Hackett et al., submitted). Importantly, this RNA expression translates into Tex19.1 protein expression (figure 3.6B). Treatment of NIH3T3 cells with Aza did not significantly affect RNA levels of *LINE1* but resulted in a ~ 5.5 fold increase of *MMERVK10C* (figure 3.6C) (two-tailed t-test, $p>0.05$ and $p<0.01$, respectively). This was surprising given the fact that no significant upregulation of *MMERVK10C* could be observed in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs compared to controls while hypomethylation was observed (figures 3.5A and B and 3.6B). I performed bisulfite sequencing which showed no de-methylation of *MMERVK10C* loci in Aza treated NIH3T3 cells compared to controls (possible explanations for this are discussed under 3.3) (figure 3.6D). Expression of the methylation dependent genes *Tex19.1* and *IAP* suggests that hypomethylation has taken place to some extent but bisulfite sequencing of those loci would be required to confirm this observation. As can be seen from figure 3.6E treatment of *Tex19.1*$^{+/+}$ and *Tex19.1*$^{-/-}$ pMEFs with Aza resulted in induction of *Tex19.1* RNA in the wild type and not the mutant cell lines and *IAP RNA* expression in wild type and mutant pMEFs as expected (two-tailed t-test, $p<0.001$ and $p<0.005$, respectively). However, the levels of de-repression were lower than observed for Aza treated NIH3T3s (figures 3.6A and 3.6E). This is likely explained by the slower replication rate in pMEFs compared to NIH3T3s. An active cell cycle is important for this experiment as cells have to replicate in order to loose methylation. Furthermore, in slowly replicating cells it takes longer for the

cytosine analogue to become incorporated into the DNA resulting in less de-methylation over a given period of time (3 days of culture in this experiment). Similarly to Aza treated NIH3T3 cells, Aza treated *Tex19.1*$^{+/+}$ and *Tex19.1*$^{-/-}$ pMEFs showed no significant change in *LINE1* RNA levels (two-tailed t-test, $p>0.05$). However, *MMERVK10C* transcript levels are significantly elevated in Aza treated cells when compared to untreated controls (two-tailed t-test, $p<0.05$) (figure 3.6E). Importantly no difference in *MMERVK10C* RNA expression between *Tex19.1*$^{+/+}$ and *Tex19.1*$^{-/-}$ pMEFs in response to Aza could be observed (two-tailed t-test, $p>0.1$) (figure 3.6E). In conclusion the here presented data suggests that upregulation of *Tex19.1* in hypomethylated fibroblasts does not influence *MMERVK10C* RNA abundance and therefore upregulation of *Tex19.1* in *p53*$^{-/-}$*Dnmt1*$^{-/-}$ cells probably does not suppress expression of *MMERVK10C.*

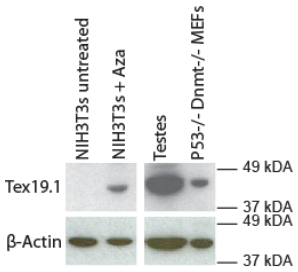
A

qPCR for *Tex19.1* and *IAP* expression in NIH3T3 cells treated with Aza and NIH3T3 untreated controls



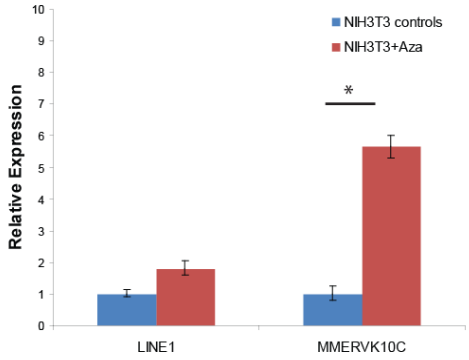
B

Western blot analysis for protein expression in NIH3T3 fibroblasts



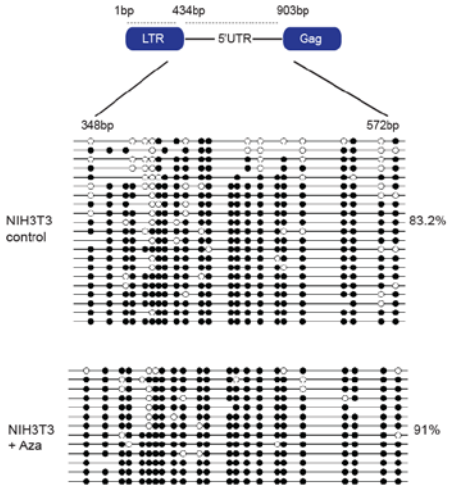
C

qPCR for *LINE1* and *MMERVK10C* expression in NIH3T3 cells treated with Aza and NIH3T3 untreated controls



D

Bisulfite sequencing of the *MMERVK10C* LTR/5'UTR region in NIH3T3 cells treated with Aza and NIH3T3 untreated controls



E

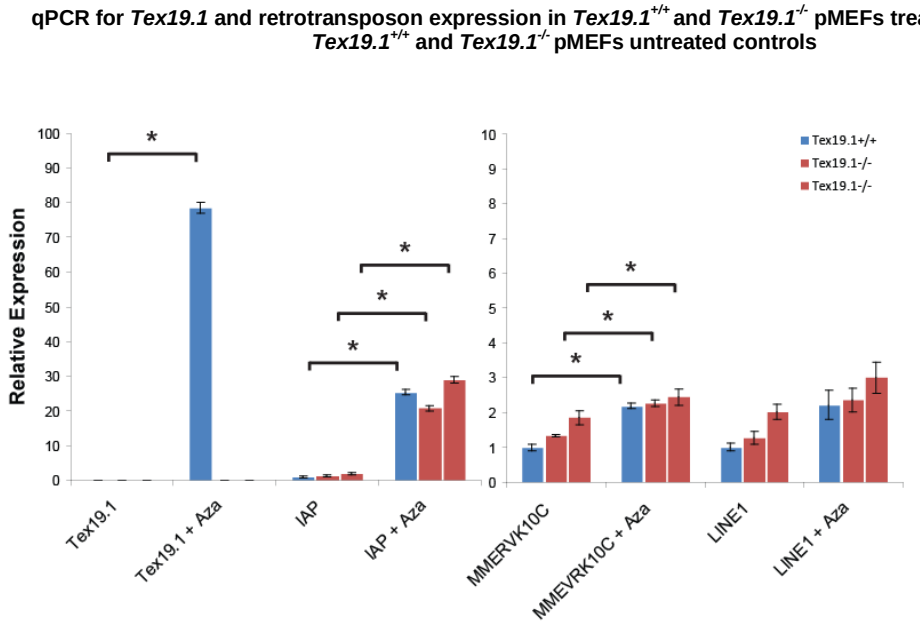


Figure 3.6 *Tex19.1* and retrotransposon expression in Aza treated NIH3T3s, *Tex19.1*$^{+/+}$ and *Tex19.1*$^{-/-}$ PEFs. (A) RNA expression levels for *IAP* and *Tex19.1* in Aza treated NIH3T3 cells compared to untreated controls (mean ± standard error for 2 biological replicates). (B) Tex19.1 protein expression in untreated and Aza treated NIH3T3 cells. Expression in testes and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ served as positive controls. (C) RNA expression levels for *LINE1* and *MMERVK10C* in Aza treated NIH3T3 cells compared to untreated controls (mean ± standard error 2 biological replicates. (D) Bisulfite sequencing showing the methylation status of *MMERVK10C* genomic loci in Aza treated NIH3T3 and control fibroblasts. Bp are shown relative to LTR start site (LTR = 1bp to 434bp; 5'UTR = 435bp to 903bp). (E) *Tex19.1*, *IAP*, *LINE1* and *MMERVK10C* RNA expression in Aza treated *Tex19.1*$^{+/+}$ and *Tex19.1*$^{-/-}$ pMEFs and untreated controls (mean ± standard error 2 biological replicates). (A, B, C and E) RNA expression levels were normalised to *β-actin* and expressed to relative controls. Asterisks indicate statistically significant differences between treated and untreated cells (two-tailed t-test, $p<0.05$).

3.2.2.2 Transcriptional regulation of *MMERVK10C* by histone modifications

Given the observation that upregulation of *MMERVK10C* RNA was observed in Aza treated NIH3T3 cells but not in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ fibroblasts I questioned if this might result from changes in histone modifications rather than DNA methylation. It has been previously reported that treatment with Aza can result in a decrease of global histone 3 lysine 9 di-methylation levels ($H3K9me^2$) (Wozniak et al., 2007). It is possible that Aza

treatment of cells results in decrease of further histone modifications but this has not been reported so far. H3K9me^2 and histone 3 lysine 9 tri-methylation (H3K9me^3) represent, together with DNA methylation and histone deacetylation, the best-characterised covalent modification associated with a repressed chromatin state (Wozniak et al., 2007; Krishnan et al., 2011). Therefore, and because repeats belonging to the ERVK subfamily have been reported to appear within chromatin regions enriched for H3K9me^3, we tested if *MMERVK10C* shows upregulation in cells with defects in establishing the H3K9 methylation mark (Day et al., 2010). Kap1 is a protein required to recruit a histone methyltransferase as well as a histone deacetylase complex to target loci (Sripathy et al., 2006). *Kap1*$^{-/-}$ ES cells show reduction in H3K9me^3 and activation of several retrotransposons including *IAP* (24-fold) and *MMERVK10* (30-fold) (Rowe et al., 2010). To test if *MMERVK10C* might also be regulated by Kap1 mediated H3K9 me^3 in somatic cells we performed qRT-PCR on *Kap1*$^{-/-}$ and *Kap1*$^{+/+}$ pMEFs for *MMERVK10C* and *IAP*. qPCR of *Kap1*$^{-/-}$ and *Kap1*$^{+/+}$ ES cells, which showed a ~14-fold and ~ 4-fold increase in *MMERVK10C* and *IAP* RNA levels, served as positive control (figure 3.7A). Conditional *Kap1*$^{-/-}$ and *Kap1*$^{+/+}$ pMEF and ES cell RNA was a kind gift from Didier Trono's laboratory. The upregulation of *IAP* elements in the ES cell sample analysed in this experiment appears to be lower compared to the published data (Rowe et al., 2010). *Kap1*$^{-/-}$ ES cells show defects in maintaining pluripotency and it is possible that in this particular experiment a large population of *Kap1*$^{-/-}$ cells has differentiated. The approximately 70% reduction in *Tex19.1* expression in *Kap1*$^{-/-}$ cells suggests that this is indeed the case (figure 3.7A). However, expression analysis of further pluripotency markers would be required to confirm this hypothesis. In contrast, *MMERVK10C* expression appears not to respond as strongly to differentiation suggesting that other or additional mechanisms achieve *IAP* silencing compared to *MMERVK10C*. *Kap1*$^{-/-}$ pMEFs show a slight increase in *IAP* expression (~1.8-fold) and possibly *MMERVK10C* expression (~1.2-fold) (figure 3.7B). Similar to the analysis performed with *p53*$^{-/-}$*Dnmt1*$^{-/-}$ and control MEFs statistical testing for significant changes of gene expression could not be performed for *Kap*$^{-/-}$ and control ES cells and pMEFs as we only had RNA for one of each cell type available. However, the low

magnitude of the changes observed in *Kap1*$^{-/-}$ pMEFs compared to controls suggests that other or additional mechanisms silence those retrotransposons in somatic cells. In the case of *IAP* elements this likely involves DNA methylation and for *MMERVK10C*, H3K9me^3 methylation might still be an important factor but is possibly facilitated in *Kap1* independent ways.

In addition to H3K9 di- and tri-methylation, histone hypoacetylation is associated with gene repression (de Ruijter et al., 2003). Histone deacetylation is achieved by histone deacetylases (HDACs) which remove acetyl groups from lysine residues and is associated with the formation of condensed and transcriptionally silenced chromatin (de Ruijter et al., 2003). In order to examine if histone deacetylation might contribute to silencing of *MMERVK10C* in somatic cells we treated NIH3T3 cells with Trichostatin A (TSA) (an HDAC class I and II inhibitor) and Aza and TSA in combination (Yoshida et al., 1990; Patel et al., 2010). Treatment of NIH3T3 cells with TSA lead to an ~ 5.6-fold increase in *Tex19.1* expression (two-tailed t-test, $p<0.05$), ~ 1.2-fold increase in *IAP* expression (two-tailed t-test, $p>0.05$) , ~ 1.3-fold increase in *LINE1* expression (two-tailed t-test, $p>0.05$) and a ~ 1.1-fold increase in *MMERVK10C* expression (two-tailed t-test, $p>0.05$). This suggests that deacetylation, mediated by HDACs that can be inhibited by TSA, is not a major contributor to *MMERVK10C* transcriptional repression in somatic cells (figure 3.7C). Upregulation of *VL30* after TSA treatment by ~ 3.8-fold (two-tailed t-test, $p<0.05$) and ~ 31-fold after Aza and TSA treatment (two-tailed t-test, $p<0.01$) suggests that inhibition of HDACs was achieved during this experiment as one particular *VL30* locus has been previously reported to show mis-expression after TSA treatment of mouse fibroblasts (Brunmeir et al., 2010). The ~ 31-fold increase of *VL30* transcript resulting from TSA and Aza treatment compared to the ~ 3.8-fold and ~ 9-fold increase resulting from TSA or Aza treatment alone suggest that for this retrotransposon DNA methylation and histone modifications act synergistically to inhibit expression. For *Tex19.1* and the retrotransposons tested, including *MMERVK10C*, expression in cells treated with Aza and TSA in combination was lower compared to cells treated with Aza alone (figure 3.7C). This suggests that for those genes the two drugs likely do not act

synergistically. Taken together the data presented here show that *Kap1* is not a major mediator of transcriptional silencing of *IAP* and *MMERVK10C* elements in somatic cells despite being of particular importance for regulating their expression in ES cells. Furthermore, silencing of *IAP* and *MMERVK10C* in somatic cells does not depend on class I and II HDACs.

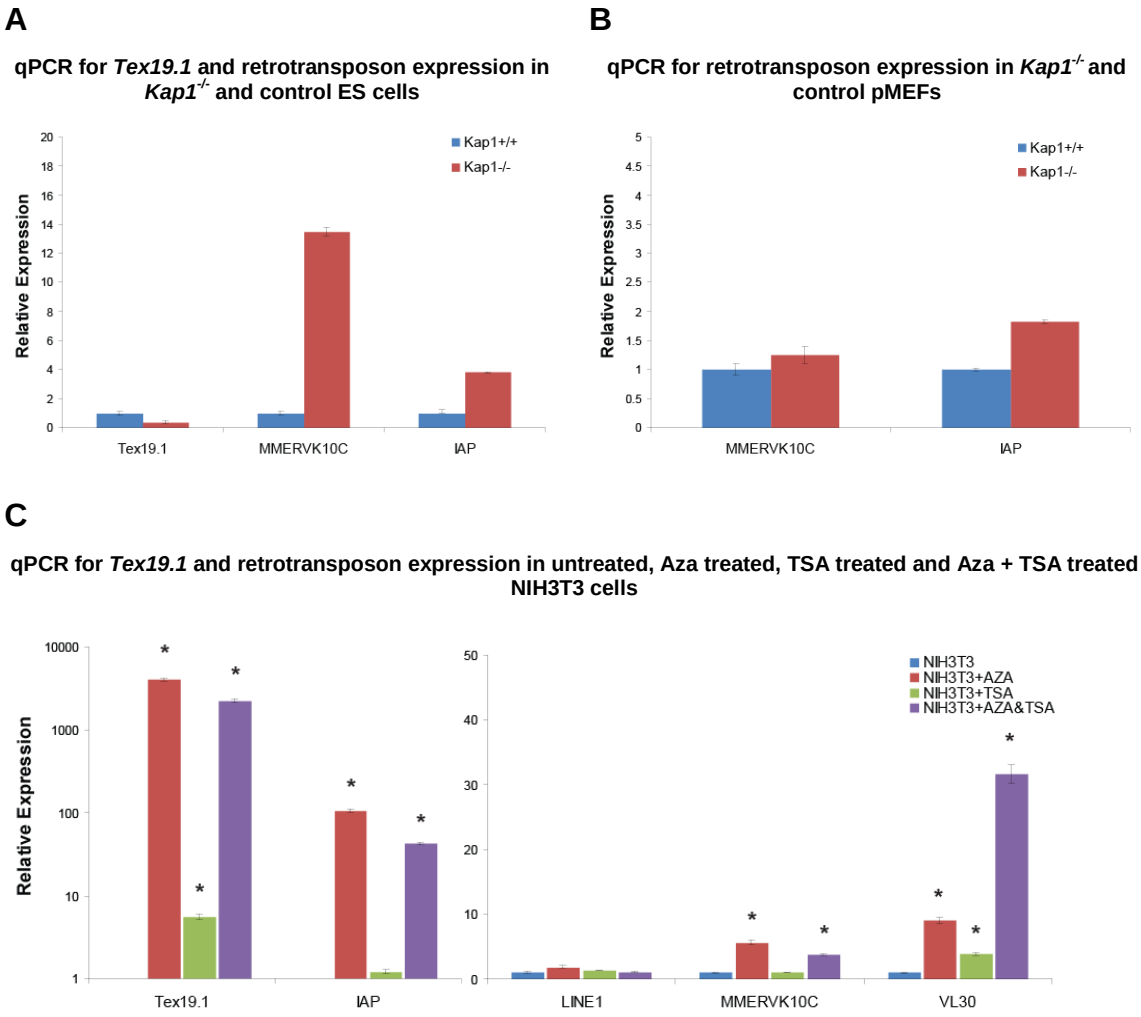


Figure 3.7 Retrotransposon expression in TSA and Aza treated NIH3T3 fibroblasts. (A) RNA expression levels *for Tex19.1*, IAP and MMERVK10C *in Kap1*$^{-/-}$ and control ES cells (mean ± standard error for 3 technical replicates). (B) RNA expression levels for MMERVK10C and IAP in *Kap1*$^{-/-}$ and control PEFs. (C) RNA expression levels for *Tex19.1* and I*AP* in Aza, TSA and Aza & TSA treated NIH3T3 cells compared to untreated controls (mean ± standard error for 3 technical replicates). (D) RNA expression levels for *LINE1*, *MMERVK10C* and *VL30* in Aza, TSA and Aza & TSA treated NIH3T3 cells compared to untreated controls (mean ± standard error 2 biological replicates. (A, B, C and D) RNA expression levels were normalised to *β-actin* and expressed relative to controls. Asterisks indicate statistically significant differences in expression between untreated and treated NIH3T3 cells (two-tailed t-test, $p<0.05$).

3.2.3 Post-transcriptional regulation of retrotransposons

Mis-expression of *MMERVK10C* RNA could be caused by failure to inhibit any stage of the retroviral life cycle. Based on Ian Adams' previous observation, that *MMERVK10C* shifts into the heavier compartment of a polysome sucrose gradient, where more actively translated mRNAs are found, it seemed possible that *Tex19.1* might inhibit *MMERVK10C* expression at the level of translation (figure 3.1).

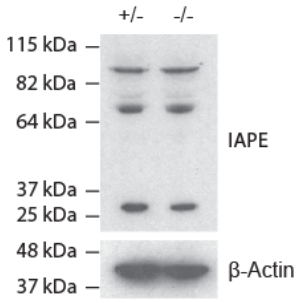
3.2.3.1 Retrotransposon protein expression in *Tex19.1*$^{-/-}$ testes

At the beginning of the project there were no antibodies available against *MMERVK10C*. We reasoned that it seems rather unlikely that *Tex19.1* should have evolved to solely repress one specific retrotransposon. Therefore, I tested other retrotransposons, for which there are established antibodies, for protein up-regulation in *Tex19.1*$^{-/-}$ testes compared to controls. First we tested the expression of *IAP* and *IAPE* elements. The retroviral families of the two elements are related but *IAPEs* contain in addition to gag and pol proteins, for assembly of intracellular viruslike particles, also a reading frame for the retroviral *env* gene (Reuss, 1992). As mentioned before *IAPE* elements (like *IAPEY3*) are closely related to *MMERVK10C* (Reichmann et al., submitted). However, as can be seen from figure 3.8A-D, *IAP* and *IAPE* do not show a change in the absence of *Tex19.1* at the protein or RNA level in 16dpp testes, suggesting that they are not targets for *Tex19.1* in male germ cells. In contrast, *LINE1* ORF1 protein (ORF1p) seems to be upregulated by about 3-fold in *Tex19.1*$^{-/-}$ 16dpp testes by western blot, as evaluated through densitometry scans, which does not appear to result from an increase of *LINE1* mRNA (figure 3.8E and F). This is in accordance with comparison of the protein levels of *LINE1* ORF1p between *Tex19.1*$^{-/-}$ and controls by immuno-fluorescence (IF) studies which showed that overall fluorescence levels significantly increase by about ~ 2-fold (two-tailed t-test, $p<0.05$) but the sub-cellular localisation of *LINE1* expression does not change between *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ testes (figure 3.9). In *Tex19.1*$^{-/-}$ mutants and controls, *LINE* ORF1p is mainly observed to be present in meiotic spermatocytes at 16

dpp. Taken together those data suggest that *Tex19.1* might affect the translation of *LINE1* mRNA or the protein stability of *LINE1* ORF1p in male germ cells.

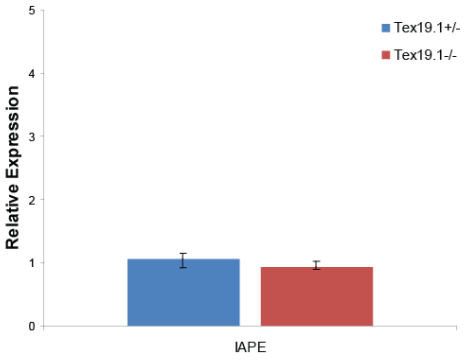
A

Western blot analysis for *IAP* protein expression in *Tex19.1*$^{-/-}$ and control testes



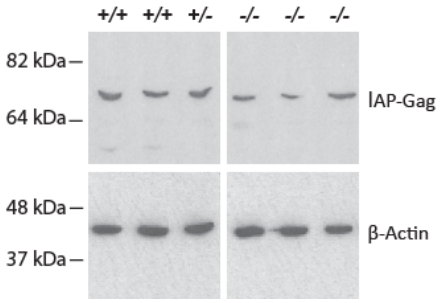
B

qPCR analysis for *IAP* RNA expression in *Tex19.1*$^{-/-}$ and control testes



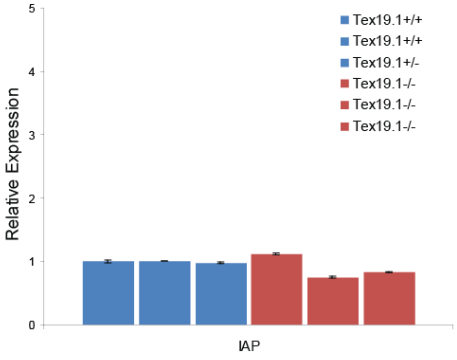
C

Western blot analysis for *IAPE* protein expression in *Tex19.1*$^{-/-}$ and control testes



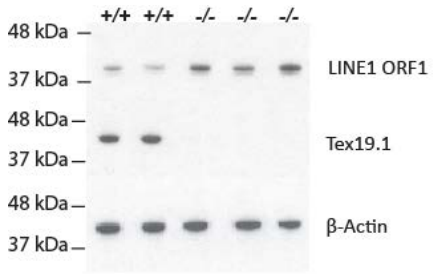
D

qPCR analysis for *IAPE* RNA expression in *Tex19.1*$^{-/-}$ and control testes



E

Western blot analysis for *LINE1 ORFp* expression in *Tex19.1*$^{-/-}$ and control testes



F

qPCR analysis for *LINE1 RNA* expression in *Tex19.1*$^{-/-}$ and control testes

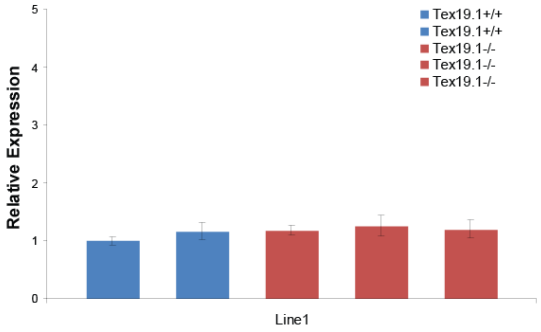
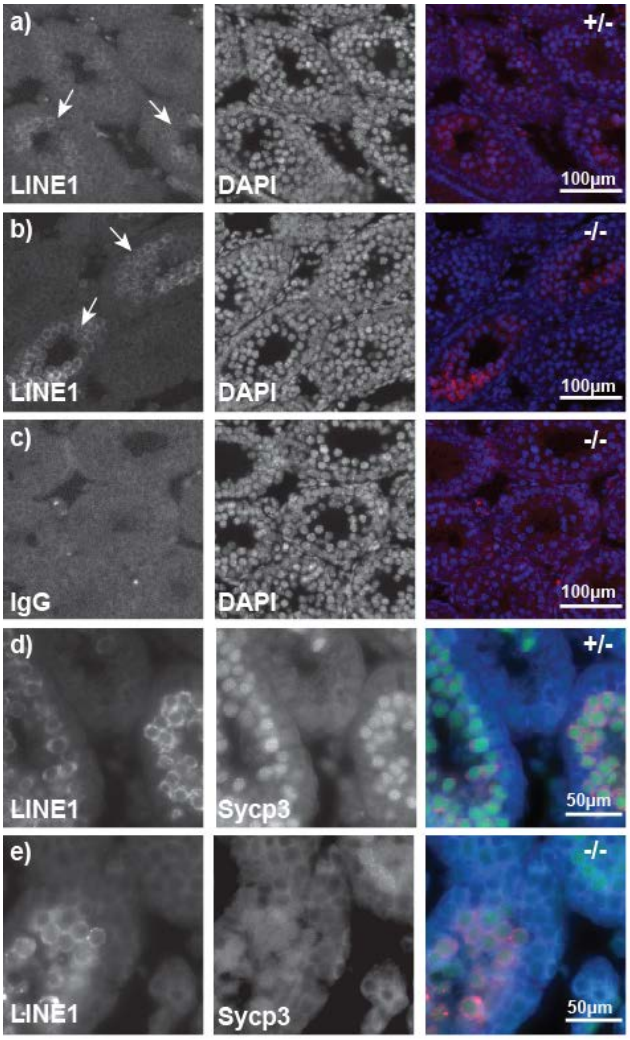


Figure 3.8 Retrotransposon protein expression in 16dpp testes (A) IAPE protein expression in 16 dpp *Tex19.1*$^{+/-}$ and $^{-/-}$ testes by western blot. (B) RNA expression levels for IAPE in 16dpp *Tex19.1*$^{+/-}$ and $^{-/-}$ testes corresponding to animals shown in A. (C) IAP protein expression in 16dpp Tex19.1+/+, *Tex19.1*$^{+/-}$ and $^{-/-}$ testes by western blot. (D) RNA expression levels for IAP in 16dpp *Tex19.1*$^{+/+}$, *Tex19.1*$^{+/-}$ and *Tex19*$^{-/-}$ testes corresponding to animals shown in C. (E) LINE1 protein expression in 16dpp *Tex19.1*$^{+/+}$, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ testes by western blot. (F) RNA expression levels for LINE1 in 16dpp *Tex19.1*$^{+/+}$, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ testes corresponding to animals shown in E. (B,D and F) graphs shown represent mean ± standard error for 3 technical replicates and RNA expression levels were normalised to *β-actin* and expressed relative to controls.

A

Immunofluorescence analysis for *LINE1* ORF1p expression in *Tex19.1*$^{-/-}$ and control testes



B

Relative *LINE1* ORF1p fluorescence in *Tex19.1*$^{-/-}$ and control testes

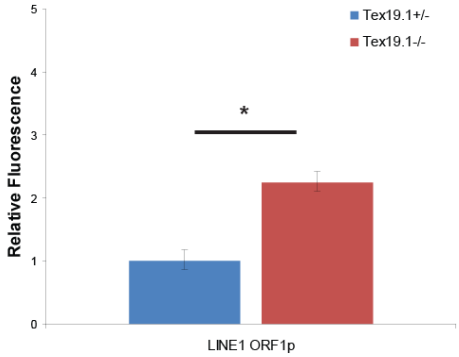


Figure 3.9 LINE1 protein expression in 16dpp *Tex19.1*$^{-/-}$ testes (A) a) IF for anti-LINE1 (red) staining in *Tex19.1*$^{+/-}$ testes. b) IF for anti-LINE1 (red) staining in *Tex19.1*$^{-/-}$ testes. c) IF for IgG (red) in *Tex19.1*$^{-/-}$ testes. d) IF for LINE1 (red) and Sycp3 (red) staining in *Tex19.1*$^{+/-}$ testes. (B) Fluorescence levels for LINE1 staining in *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ testes (mean ± standard error for three animals) was normalised to IgG and expressed relative to littermate controls. Asterisks indicate statistically significant differences (two-tailed t-test, $p<0.05$).

3.2.3.2 Retrotransposon protein expression in *Tex19.1*$^{-/-}$ ES cells

In order to test if translational regulation of retrotransposons might be affected in *Tex19.1*$^{-/-}$ mutant ES cells we performed western blots for IAP, IAPE and LINE1 protein expression in *Tex19.1*$^{-/-}$ ES cells. IAP protein expression appears relatively uniform between the ES cells lines used (figure 3.10A). In contrast *IAP* RNA expression in *Tex19.1*$^{-/-}$ and control ES cells seems more variable than in testes but does not segregate with protein abundance or genotype (figure 3.10B). This is also largely the case for *IAPE* RNA and protein expression for the cell lines studied, however, one *Tex19.1*$^{-/-}$ cell line behaves as outlier and shows an about 25-fold upregulation of *IAPE* RNA compared to controls (figure 3.10D). This increase in RNA is not reflected by an increase in protein (figure 3.12C). *LINE1* RNA expression appears to be variable between all cell lines analysed and the same is true for *LINE1* protein expression in male and female ES cells (figure 3.10E and F). In contrast to the data presented in figure 3.9E, which demonstrated upregulation of *LINE1* ORF1p in *Tex19.1*$^{-/-}$ testes compared to controls, *LINE1* ORFp appears to be highly variable between different ES cell lines. The discrepancy between the two data sets could be caused by the fact that ES cells might generally show high variation of LINE ORF1p expression levels between cell lines which would make it difficult to pick up small changes in protein abundance. Alternatively, *LINE1* might not represent a target for Tex19.1 in ES cells.

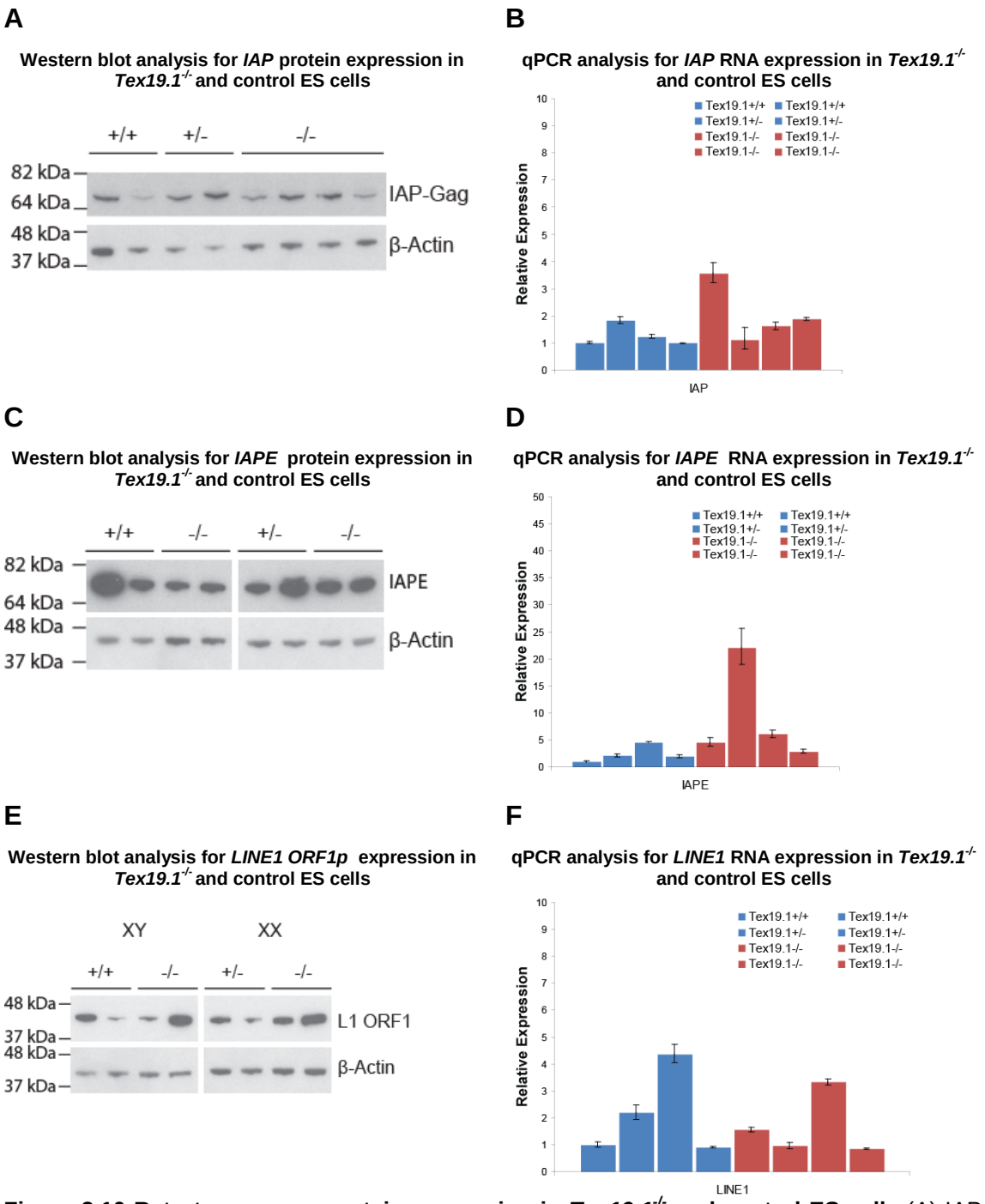


Figure 3.10 Retrotransposon protein expression in *Tex19.1*$^{-/-}$ and control ES cells (A) IAP protein expression in *Tex19.1*$^{-/-}$ and control ES cells by western blot. (B) RNA expression levels for IAP in *Tex19.1*$^{-/-}$ and control ES cells corresponding to cell lines shown in A. (C) IAPE protein expression in *Tex19.1*$^{-/-}$ and control ES cells by western blot. (D) RNA expression levels for IAPE in *Tex19.1*$^{-/-}$ and control ES cells corresponding to cell lines shown in C. (E) LINE1 protein expression in *Tex19.1*$^{-/-}$ and control ES cells lines by western blot. (F) RNA expression levels for LINE1 in *Tex19.1*$^{-/-}$ and control ES cells lines shown in E. (B,D and F) Graphs represent mean ± standard error for 3 technical replicates.

In order to directly test if Tex19.1 might regulate *LINE1* translation in ES cells I transfected *Tex19.1*$^{-/-}$ and control ES cells with reporter plasmids containing *LINE1* sequence. The plasmids were kindly provided by Sandy Martin from the University of Colorado. Sandy Martin's laboratory has previously reported that sequences upstream of ORF1 and ORF2 are involved in translational regulation of *LINE1* elements (Li et al., 2006). *LINE1* RNA is dicistronic when its two ORFs are translated and it has been suggested that mechanisms, alternative to the standard cap-dependent recognition system which is followed by ribosome scanning to the first AUG, are needed for efficient translation of both proteins (Li et al., 2006). One of the mechanisms that allow translation of the RNAs in a cap-independent manner is via an internal ribosomal entry side (IRES). IRESes are found in a wide variety of viral and cellular RNAs and functionally defined by their ability to promote independent translation of the second cistron in a dicistronic RNA. Using this definition, it was found that the sequences upstream of ORF1 and ORF2 in mouse *LINE1* RNA are IRESes (Li et al., 2006). In order to test if Tex19.1 might regulate *LINE1* translation via its 5'UTR, in particular via the IRES, we transfected *Tex19.1*$^{-/-}$ and control ES cells with reporter plasmids containing *LINE1* sequence. The vector backbone of the constructs used here was pRF, a dicistronic reporter vector that contains renilla luciferase (rluc) cloned downstream of a composite SV40/T7 promoter as the first cistron, and of firefly luciferase (fluc) as the second cistron. A highly structured element, between the two reporter genes but upstream of the cloning site for the LINE1 sequences, serves as a barrier to expression of fluc by read-through translation (figure 3.11A). Fluc expression is therefore indicative of IRES activity. The cells were transfected with the empty vector (pRF) and a vector containing parts of the *LINE1* 3'UTR (pRF3) which both served as negative controls (figure 3.11B). The known CrPV intergenic IRES served as a positive control (pRFCrPV) (figure 3.11B). As expected *Tex19.1*$^{-/-}$ and control ES cells showed an increase pRFCrPV fluc expression by ~2.6 to 8-fold compared to the negative controls (figure 3.11C). For pRFD and pRFA, which contain 400 nt upstream of the first *LINE1* AUG and a *LINE1* fragment extending 201 nt upstream of the ORF2 AUG respectively, it was found that those sequences promote translation over empty vector and negative

control but no difference between *Tex19.1*$^{-/-}$ and control ES cells could be observed (figure 3.11B and C). It is worth mentioning that three active families of *LINE1* elements (Tf-, Gf-, and A-type) can be distinguished in the mouse genome based on different repeated 200 bp monomer units within their 5' UTRs (Ostertag and Kazazian, 2001). The element tested in this experiment belongs to the Tf-type and it can not be excluded that Tex19.1 might target Gf- and/or A-type *LINE1s* in ES cells and male germ cells but not Tf-type *LINE1s* (Li et al., 2006). However, it is also possible that *LINE1* is not a translational target for Tex19.1 in ES cells.

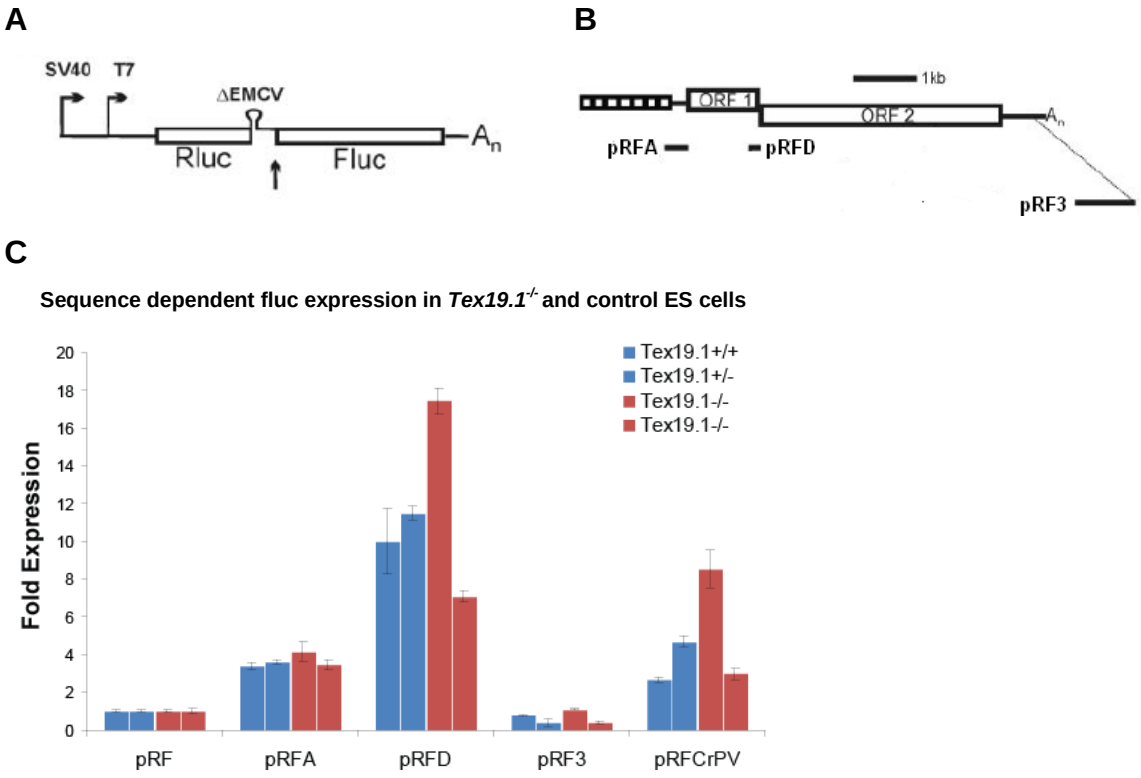


Figure 3.11 LINE1 5'UTR dependent fluc expression in *Tex19.1*$^{-/-}$ and control ES cells. (A) Diagram of the pRF dicistronic reporter vector containing firefly (Fluc) and Renilla (Rluc) luciferase genes. The dicistronic mRNA can be transcribed by an SV40 promoter/enhancer sequence in vivo or by a T7 promoter in vitro. ΔEMCV between the two reporter genes upstream of the cloning site for the LINE1 sequences serves as a barrier to expression of fluc by read-through translation. (B) Diagram of a LINE1 element. LINE1 sequences cloned into pRF dicistronic reporter reporter are indicated pRFA, pRFD and pRF3. (C) Relative luciferase activities (to pRF ± standard error) of the 400nt (pRFD) or 201nt (pRFA) upstream of ORF1 and ORF2, pRF3 (LINE1 3'UTR) and pRFCrPV in *Tex19.1*$^{-/-}$ and control ES cells. Diagrams presented in A and B were taken and modified from Li et al., (2006).

3.2.3.3 Does Tex19.1 act as a translational regulator of retrotransposons?

LINE1 RNA and *LINE1* ORF1p appeared variable in *Tex19.1*$^{-/-}$ and control ES cells. Together with the luciferase assay those data suggest that Tex19.1 is not a translational regulator of *LINE1* in ES cells. However, the increase of *LINE1* protein in *Tex19.1*$^{-/-}$ testes, the cytoplasmic localization of the Tex19.1 protein as well the northern blot analysis, prepared by Ian Adams, shown in figure 3.1, suggest that Tex19.1 might act as a translational regulator of *LINE1* and *MMERVK10C* in male germ cells. In order to test if Tex19.1 might inhibit translation by direct binding of target mRNAs in testes I used oligo (dT) cellulose in order to isolate polyadenylated mRNAs from 16 dpp wild type testes extract. Proteins associated with mRNAs were eluted from the beads by boiling in Laemmli buffer and then analyzed by western blot. Figure 3.12 shows that Pabp1, which served as a positive control, can be captured by binding polyadenylated RNAs to the oligo (dT) cellulose. This seems to be specific to polyadenylated RNAs as blocking of the oligo dT beads with poly A does largely abolish Pabp1 binding to the beads. In contrast Tex19.1 protein is clearly detectable in the input but can not be found in the sample eluted from the oligo (dT) cellulose. This suggests that Tex19.1 is not associated with polyadenylated RNAs under the native pull down conditions used in this experiment.

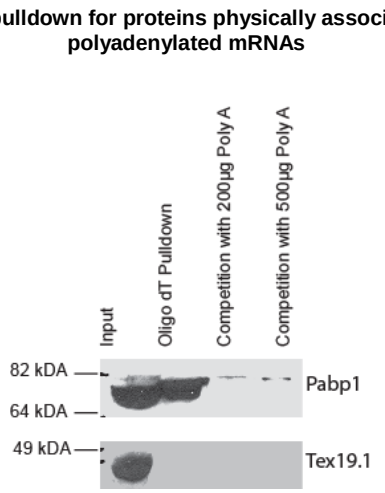


Figure 3.12 Tex19.1 is not associated with polyadenlyated mRNAs in mouse testes. Western blot for 16 dpp testes extract. Extract was subjected to binding by oligo(dT) cellulose and proteins bound to mRNA were eluted from the beads. Competition with poly A chains served as control. Samples were blotted for Pabp1 and Tex19.1.

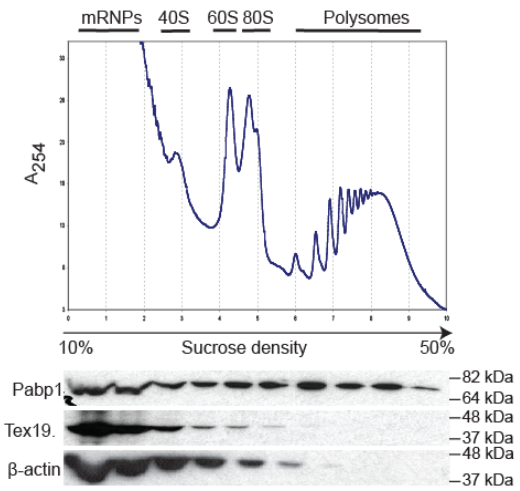
In order to analyse if the Tex19.1 protein is associated with polysomes and therefore possibly the translational machinery I performed sucrose gradients of 16 dpp wild type testes and ES cells extract. In order to do so cells or tissues were lysed in the presence of cycloheximide, which acts to block peptidyl transfer and prevent ribosome run off, or RNAse which causes disaggregation of the ribosomes (Fenwick, 1968). Lysates were then centrifuged through 10-50% sucrose gradients and absorbance at 254 nm was used to establish the position of polysomes, monosomes (80S ribosomes), ribosomal subunits and mRNPs. Fractions from the gradient were methanol precipitated and western blotted for Tex19.1, Pabp1 and β-actin. Pabp1, which served as a positive control, was detected throughout the gradient, including the polysomal fractions in the presence of cycloheximide treatment and relocalised to lighter fractions upon RNAse treatment in ES cells and testes as expected (figure 3.13A, B, C and D) (Sachs and Davis, 1989; Gray et al., 2000). β-actin, which is not RNA associated, can be found in the mRNP and monosome fractions but is not detectable in the heavier compartment of the sucrose gradient and its localisation does not change upon RNAse treatment (figure 3.13A, B, C and D). Figure 3.13A and C show that Tex19.1 protein is predominately found in the mRNP fractions of a sucrose polysome gradient in wild type 16 dpp testes and wild type

ES cells. This localization does not seem to be RNA dependent as Tex19.1 protein does not shift into lighter fractions if samples are treated with RNAse before being analysed on the sucrose gradient (figure 3.13B and D). The slight shift in Tex19.1 protein seen in figure 3.13B likely is not caused by the RNAse treatment but by small difference in the gradient itself and/or efficiency of protein precipitation prior to western blot analysis. This is based on the observation that comparison of Tex19.1 protein localisation on a sucrose gradient from several animals did not result in a consistent shift of Tex19.1 protein in response to RNAse treatment (not shown). Similarly, treatment of ES cell extract with EDTA (EDTA chelates Mg^{2+} ions which causes protein complexes to dissociate) and with EDTA and RNAse in combination showed no consistent shift of Tex19.1 protein into lighter fractions of a 10-50% sucrose polysome gradient. However, Pabp1, which served as control, did shift upon treatment with EDTA or EDTA and RNAse into the lighter fractions of the sucrose gradient in ES cells. It is possible that, since Tex19.1 protein does migrate in the lighter fractions of the polysome gradient to start with, the system is not sensitive enough to see a possible shift upon RNAse and/or EDTA treatment. Based on the data presented here it is reasonable to conclude that Tex19.1 does not localise with the actively translated part of the sucrose gradient and hence it is plausible that Tex19.1 might directly or indirectly prevent access of ribosomes or other components of the translational machinery to mRNAs, i.e. *MMERVK10C*. This could explain the shift of *MMERVK10C* mRNA into the more actively translated part of the sucrose gradient seen in *Tex19.1*$^{-/-}$ testes (figure 3.1). In order to verify the data shown in figure 3.1, as well as to analyze if the increase of *LINE1* protein might be the result from an increase in translation in the absence of Tex19.1, we performed polysome gradients from *Tex19.1*$^{-/-}$ and control 18dpp testes. The difference in the polysome profiles between figure 3.13 and 3.14 results from the fact that different gradient stations and therewith different detection and fraction collection systems were used for the two experiments. To analyse the abundance of RNAs in different compartments of the sucrose gradient the fractions classified as mRNPs and 40S peaks, 60S and 80S as well as polysomes were pooled and RNA precipitated with Trizol LS reagent (figure 3.14A). For simplicity pooled mRNP and

40S fractions are referred to as mRNPs, 60S and 80S as monosomes and the remaining fractions as polysomes. I performed qRT-PCR for *LINE1* and *MMERVK10C* RNA. As can be seen from figure 3.14B and C *MMERVK10C* expression was between ~2.5 to 4-fold upregulated in *Tex19.1*$^{-/-}$ testes input fractions whereas *LINE1* expression appeared similar between *Tex19.1*$^{-/-}$ and control animals. This is consistent with the data presented in figures 3.4 and 3.8. The majority of the over-expressed *MMERVK10C* RNA was found to be localized to the mRNP fractions and no shift of *MMERVK10C* transcript into the monosome or polysome compartment could be observed in the absence of *Tex19.1* using this assay. Similarly, *LINE1* transcript was not enriched in the monosome or polysome compartment in *Tex19.1*$^{-/-}$ mutants compared to controls. In summary, I have shown here that Tex19.1 appears not to be associated with polyadenylated RNAs under the native pull down conditions used. Moreover, Tex19.1 does not localise to polysomes in wild type testes and ES cells, the localisation of Tex19.1 on a sucrose gradient profile does not change upon RNAse or EDTA treatment. In addition *LINE1* and *MMERVK10C* transcripts do not shift into the heavier compartment of the sucrose gradient in *Tex19.1*$^{-/-}$ mutant testes. In conclusion the here presented experiments, together with the luciferase experiment presented in figure 3.11, suggest that Tex19.1 is not a translational regulator of *LINE1* or *MMERVK10C*. These observations and the discrepancy between the data presented in figure 3.1 and 3.14 are discussed under 3.3.

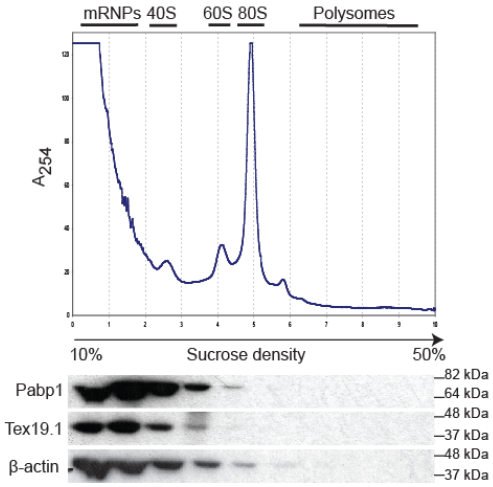
A

Sucrose gradient from 16dpp wild type testes treated with cycloheximide



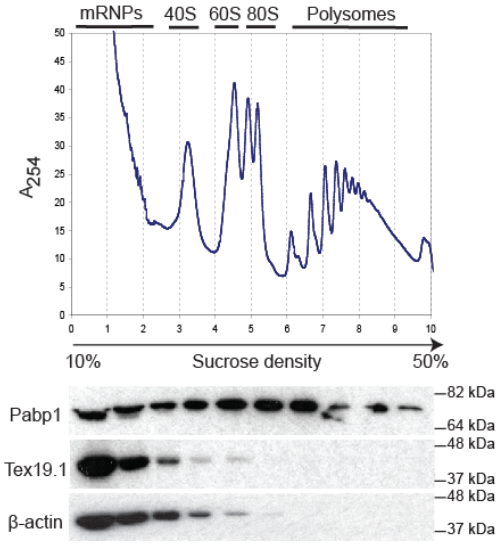
B

Sucrose gradient from 16dpp wild type testes treated with RNAse



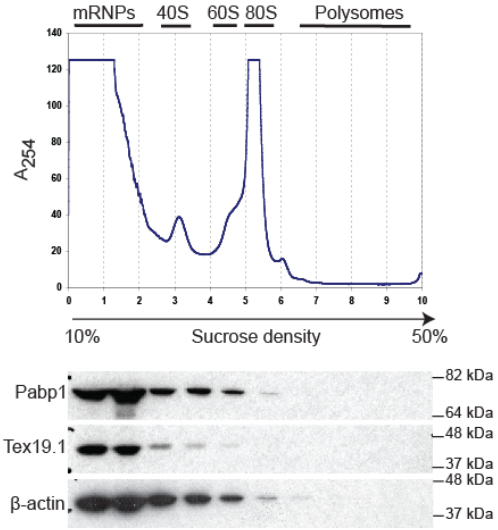
C

Sucrose gradient from wild type ES cells treated with cycloheximide



D

Sucrose gradient from wild type ES cells treated with RNAse



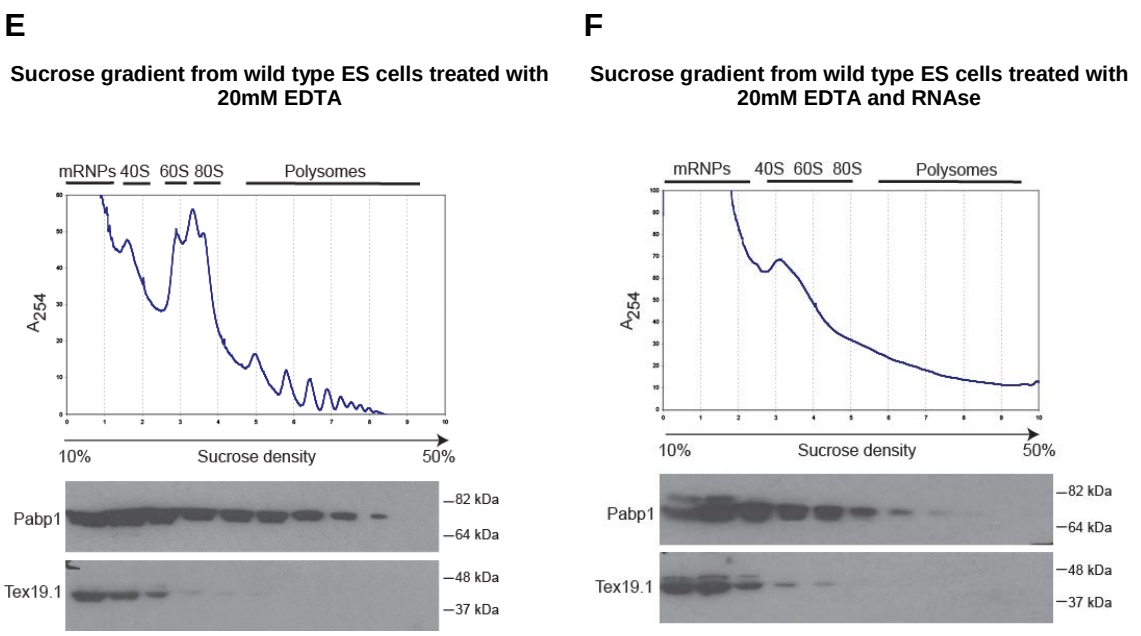


Figure 3.13 Sucrose gradient on 16dpp wild type testes and wild type ES cells. (A and B) 16dpp testes extract was treated with cyclohexamide (A) or cyclohexamide and RNAse (B) and individual fractions blotted for Pabp1, Tex19.1 and β-actin. (C and D) Wild type ES cells extract was treated with cyclohexamide (C) or cyclohexamide and RNAse (D) and individual fractions blotted for Pabp1, Tex19.1 and β-Actin. (E and F) ES cells extract was treated with EDTA (E) or EDTA and RNAse (F) and individual fractions blotted against Pabp1 and Tex19.

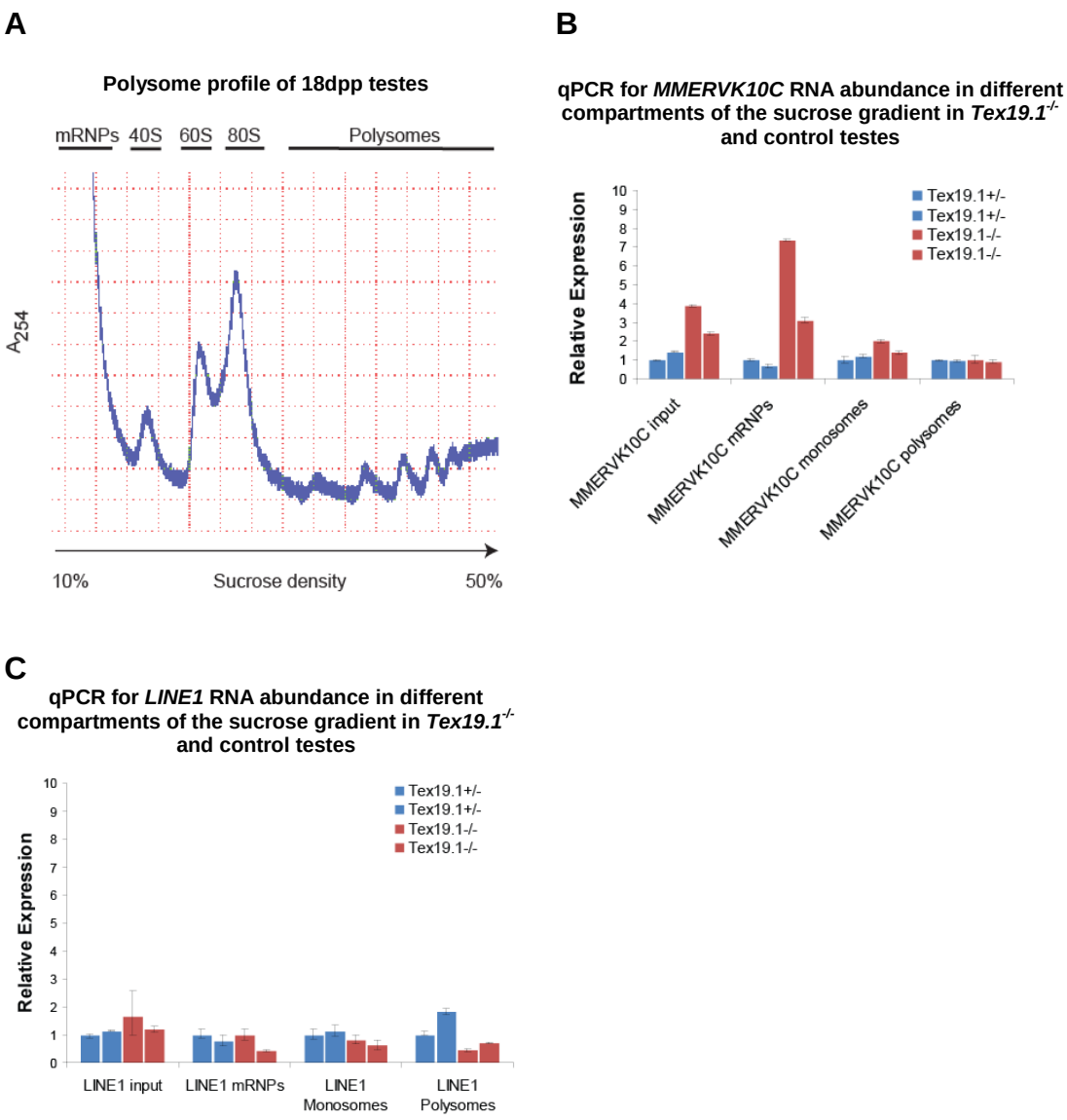


Figure 3.14 qRT-PCR on sucrose gradient fractions from 18dpp *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ testes. (A) Sucrose gradient profile for 18dpp testes showing fractions pooled for each category. (B) Transcript levels for MMERVK10C (mean ± standard error for technical replicates) in *Tex19.1*$^{-/-}$ testes were normalized to *β-actin* and expressed relative to one of the littermate controls. (C) Transcript levels for LINE1 (mean ± standard error for technical replicates) in *Tex19.1*$^{-/-}$ testes were normalized to *β-actin* and expressed relative to one of the littermate controls.

3.2.4 What are the interaction partners of Tex19.1?

The data presented in the previous section show that Tex19.1 does not move into lighter fractions of a polysome gradient upon RNAse treatment and does not strongly associate with polyadenylated mRNAs in testes extract. In order to gain a better understanding how Tex19.1 functions during embryonic and germline development and retrotransposons repression we examined with which proteins Tex19.1 might interact. In order to investigate if Tex19.1 protein is part of a complex we performed a sucrose size gradient. Figure 3.15A shows that the majority of Tex19.1 protein in cytoplasmic ES cell extract migrates at a size of 150 – 443 kDa within the sucrose gradient. No Tex19.1 protein migrating within the sucrose at 44kDa, the predicted molecular weight of Tex19.1, could be observed by western blot. This suggests that Tex19.1 might migrate at a different size than its predicted molecular weight due its native confirmation or because the majority of detectable Tex19.1 protein is part of a complex. Ian Adams was able to show by mass-spectrometry of GFP pull-downs from 293 cells stably transfected with GFP-tagged Tex19.1 that Tex19.1 interacts with the E3 ubiquitin ligase UBR2. This is in accordance with a recent study which identified an interaction between Tex19.1 and UBR2/Ubr2 in HeLa and male germ cells (Yang et al., 2010). During the course of this thesis I was able to confirm by western blot that Tex19.1 also interacts with Ubr2 in ES cells (pull-down was performed by Chao-Chun Hung) (figure 3.16B).

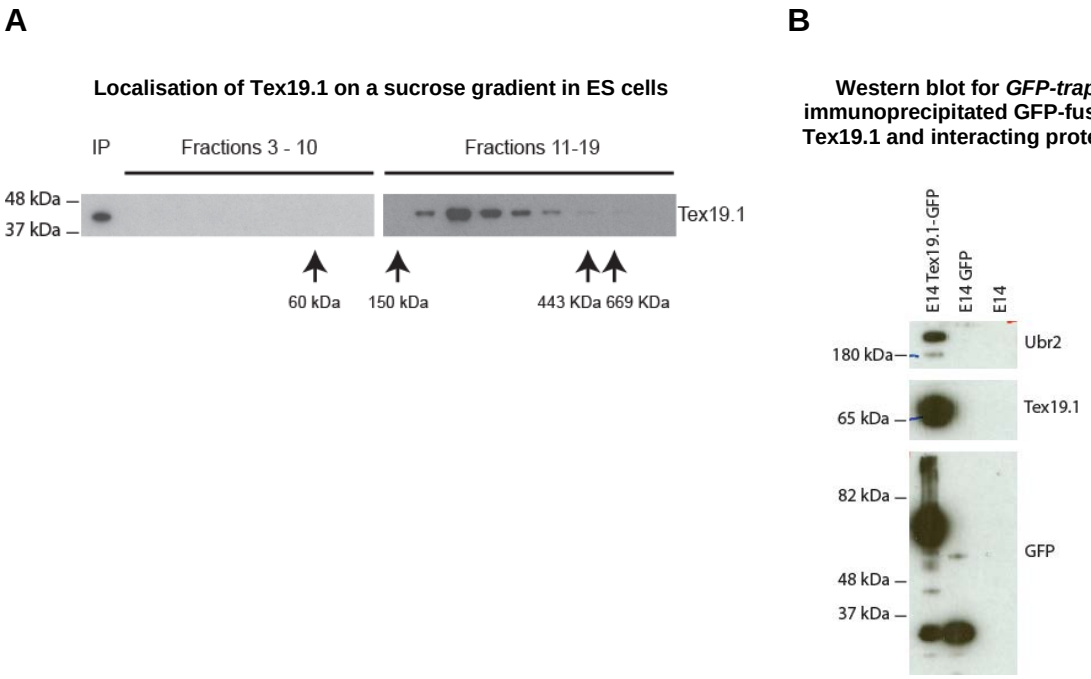


Figure 3.15 Tex19.1 interacts with Ubr2 in 293 cells and ES cells. (A) 293T cells and 293T cells stably expressing Tex19.1-GFP were subjected to GFP-TRAP immuno-precipitation. Pellets were analyzed by western blot with the indicated antibodies. (B) ES cells, ES cells stably expressing GFP and ES cells stably expressing Tex19.1-GFP were subjected to GFP-TRAP immuno-precipitation. Pellets were analyzed by western blot with the indicated antibodies.

This prompted me to examine if Tex19.1 and Ubr2 might function together in the same pathway and if this interaction could, at least in part, explain the *Tex19.1*$^{-/-}$ phenotypes. To address if Tex19.1 influences Ubr2 function and acts as a component of the N-end rule pathway of protein degradation, we used a reporter system that uses green fluorescent protein (GFP)-based substrates for the quantification of ubiquitin/proteasome-dependent proteolysis in living cells (Dantuma et al., 2000). The expression plasmids used for this reporter system use N-end rule and ubiquitin fusion degradation (UFD) signals to convert the stable jellyfish GFP into a substrate for ubiquitin–proteasome-dependent proteolysis (Dantuma et al., 2000). In this study I used four Ub-GFP plasmids, one negative and one positive control and two N-end rule substrates (figure 3.16A). The negative control Ub-M-GFP does not contain a degradation signal upon ubiquitin cleavage and is therefore expected to be as stable as unmodified GFP. The UFD substrate Ub-G76V-GFP or Ub-G-GFP, which served as

positive control, contains a mutated uncleavable ubiquitin moiety and should therefore be constantly targeted for degradation. The two plasmids that serve as N-end rule substrates contain arginine (Ub-R-GFP) and leucine (Ub-L-GFP) as destabilizing amino acids with Ub-R-GFP being more efficiently targeted than Ub-L-GFP. Firstly, we tested if all 4 GFP plasmids would show the expected expression pattern when transfected into HeLa cells. Cells were transfected with the GFP plasmids in combination with a far-red expression plasmid, which served as a control for transfection efficiency, and FACS analysed 48h post-transfection. Figure 3.16B shows that transfection of the four plasmids into HeLa cells gave the expected result with Ub-M-GFP being the most stable plasmid of the four, Ub-R-GFP and Ub-L-GFP being less stable and Ub-G-GFP being the least stable. We then transfected those plasmids into two *Tex19.1*$^{-/-}$ ES cell lines and two control lines and observed no Tex19.1 dependent difference between the GFP expression levels of either plasmid (figure 3.16C). However, when we transfected the plasmids into HeLa cells with an empty vector and a vector expressing Tex19.1 protein we observed a statistically significant reduction in the amount of Ub-GFP-L expression in cells transfected with Tex19.1 compared to the empty vector control suggesting an increase in N-end rule pathway activity (two-tailed t-test, $p<0.05$). Possibly reasons for the discrepancy between loss of function experiment in the ES cells and the gain of function experiment in the HeLa cells are discussed in 3.3. Taken together the data presented here suggest that Tex19.1 interacts with Ubr2 in ES cells and that Tex19.1 might promote activity of N-end rule mediated proteolysis.

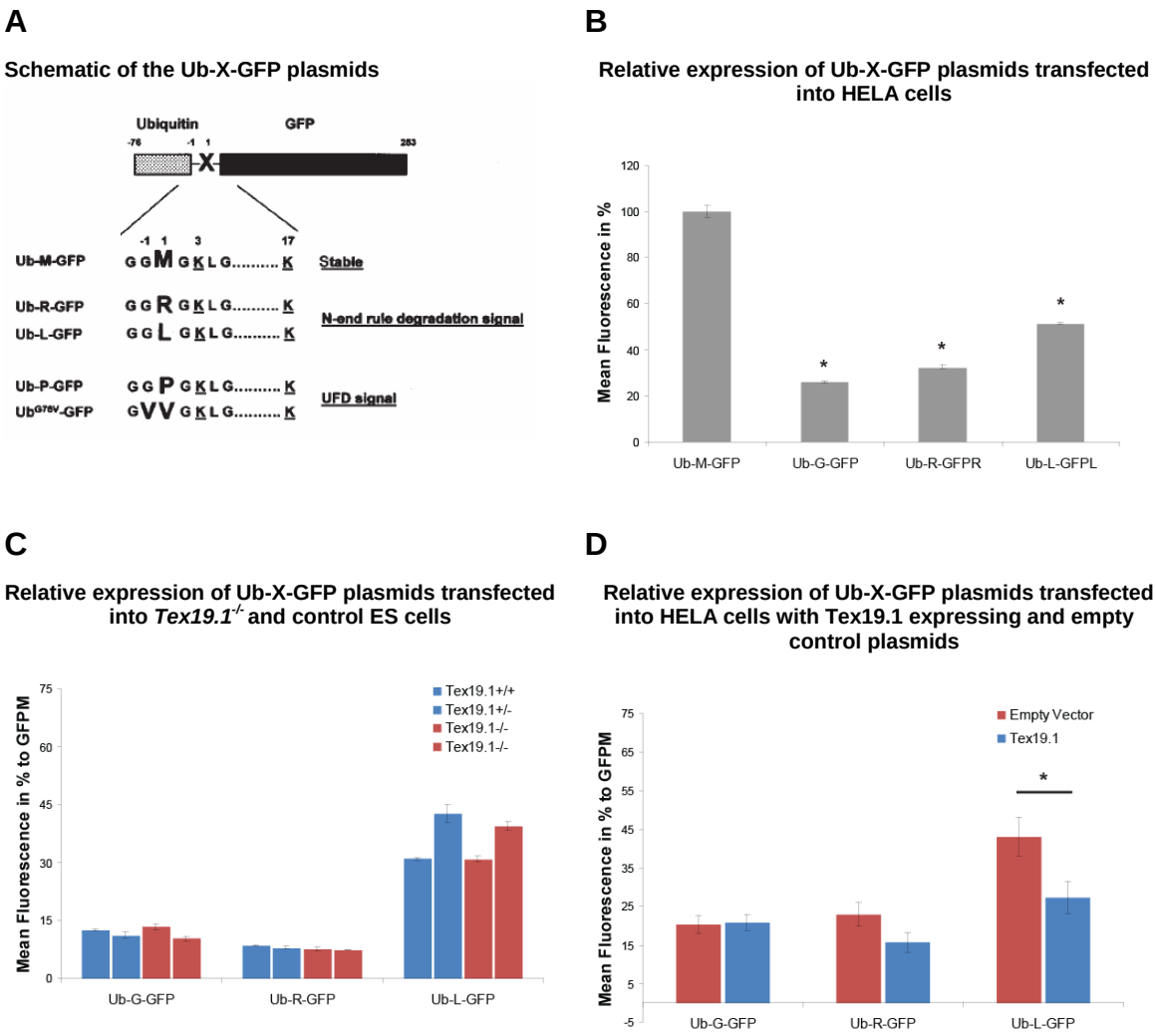


Figure 3.16 Tex19.1 dependent Ub-X-GFP expression monitored by fluorescence intensity. (A) Diagram of the Ub-X-GFP plasmids. The amino acid in bold represents the N-degron (R and L), the UFD signal (Ub^{G76V}-GFP will be referred to as Ub-G-GFP) or the residue referring stability (M). The lysine residues in position 3 and 17 are potential ubiquitination sites, taken from Dantuma et al., (2000). (B-E) Flow-cytometric analysis of transiently transfected HeLa (B and D) or *Tex19.1*$^{-/-}$ and control ES cells (C) expressing the Ub-X-GFP plasmids. GFP expression levels were normalised to far-red expression levels and Ub-M-GFP. Asterisks indicates statistically significant change (two-tailed t-test, $p<0.05$). (D) HeLa cells were transfected with Ub-X-GFP plasmids in combination with empty and Tex19.1 expression vector. (B and C) or mean ± standard error for three biological replicates (D and E).

3.3 Discussion

In this chapter I hoped to better understand the role of *Tex19.1* in retrotransposon suppression. Firstly, I verified the repertoire of repetitive elements upregulated in the absence of *Tex19.*1 in the testes which were identified using a novel, repeat-annotation method of microarray data (this was performed by Ian Adams). Next, I set out to learn more about how the relatively obscure *MMERVK10C* retrotransposon is transcriptionally regulated. I also tested the hypothesis that *Tex19.1* affects translational efficiency of retrotransposons. I finished the chapter by investigating the interaction of Tex19.1 with Ubr2 and I provide preliminary evidence that Tex19.1 modulates N-end rule pathway degradation of target proteins.

3.3.1 Retrotransposon RNA upregulation in *Tex19.1*$^{-/-}$ males

The extent of retrotransposon de-repression in *Tex19.1*$^{-/-}$ testes was investigated by taking advantage of repeat-annotated Illumina Mouse WG-6 v2.0 Beadchip probes. Using this method we found that *MMERVK10C*, but not related retrotransposons, are upregulated in *Tex19.1*$^{-/-}$ testes in inbred C57BL/6 mice at 16dpp which we were able to confirm by qRT-PCR (figure 3.4). The up-regulation of *MMERVK10C* is consistent with the previously reported over-expression of this element in *Tex19.1*$^{-/-}$ animals on a mixed (129/Ola x CD1) genetic background (Öllinger et al., 2008). Interestingly Mary Taggart, in Ian Adams' laboratory, was able to show that the upregulation of *MMERVK10C* RNA is caused by mis-expression of multiple genomic copies suggesting that de-repression does not occur at a single locus. It is not clear why we see RNA upregulation of *MMERVK10C* but not i.e. *IAPs*, *LINEs* or *SINEs* in *Tex19.1*$^{-/-}$ mutants. One of the questions to answer is where this specificity is coming from, and if mouse mutants in i.e. *Dnmt3L*, *Miwi2* and *Mili*, all of which also have defects in meiotic progression correlating with defects in retrotransposon repression, show *MMERVK10C* upregulation or if the de-repression of this particular element is restricted to *Tex19.1*$^{-/-}$ germ cells. In

those mutants data about *MMERVK10C* have not been published and instead only the better characterized *IAP* and *LINE1* elements were examined (Bourc'his and Bestor, 2004; Kuramochi-Miyagawa et al., 2004; Carmell et al., 2007). Michael Ling from Hamish Scott's laboratory analysed *MMERVK10C* expression in 8dpp *Dnmt3L*$^{-/-}$ and control testes and showed that there is no significant difference in expression between mutants and controls (Michael Ling, personal communication). This suggests that upregulation of *MMERVK10C* is not a general result of defects in meiotic chromosome synapsis and that *Dnmt3L* is dispensable for *MMERVK10C* silencing in male germ cells. On the contrary, *Tex19.1* silences *MMERVK10C* RNA expression in the testes but appears not to be required for silencing of *IAP*, *LINE1* or even elements closely related to *MMERVK10C*. This and the microarray findings presented here suggest that different retrotransposons and even closely related elements differ in sensitivity to regulatory mechanisms in developing germ cells and that *Tex19.1* specifically promotes silencing of *MMERVK10C* expression during spermatogenesis. However, the basis for this specificity is still unclear.

3.3.2 Transcriptional regulation of *MMERVK10C* by DNA methylation

As summarized under 1.7.2.2 in the majority of mouse mutants that upregulate retrotransposons and show defects in the progression of meiosis in the male germ line, retrotransposon de-repression is thought to be ultimately mediated by defects in DNA methylation which leads to a loss of transcriptional suppression (Bourc'his and Bestor, 2004; Kuramochi-Miyagawa et al., 2004; De La Fuente et al., 2006; Carmell et al., 2007). Analysis of *MMERVK10C* expression in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ fibroblast showed no upregulation despite being 50% demethylated compared to controls. This suggests that *MMERVK10C* behaves different from the well characterised *IAP* elements in somatic cells as *IAPs* are strongly upregulated in hypomethylated fibroblasts (figure 3.5). Interestingly, *LINE1* RNA also showed no significant change in abundance in those cells (figure 3.5). In ES cells silencing of *LINE1* transcription does not seem to be dependent

on DNA methylation as ES cells null for *Dnmt1* and the *de novo* methyltransferases do not exhibit a significantly change in *LINE1* expression levels (Reichmann et al., in submission). However, in *Dnmt3L* mutant testes, where *de novo* methylation of retrotransposons is impaired, *LINE1* shows upregulation at the RNA level by northern blot (Bourc'his and Bestor, 2004). Similarly, *Miwi2* and *Mili* mutant germ cells upregulate *LINE1* RNA which is also correlating with loss of DNA methylation at the promoter (Kuramochi-Miyagawa et al., 2008). *MeCP2*$^{-/-}$ mutant neurons, which are null for a protein which is thought to be involved in DNA methylation recognition, also show elevation of *LINE1* transcription (Muotri et al., 2010). Furthermore, a study performed by Jackson-Grusby et al., (2001) reports an about 3-fold upregulation of *LINE1* RNA in *Dnmt1*$^{-/-}$ mutant fibroblasts suggesting that the *LINE1* promoter is able to drive expression in fibroblasts. The discrepancy in the *LINE1* expression levels presented in this thesis and the study by Jackson-Grusby et al., (2001) could result from differences in the cell lines used for the analysis. The fibroblasts used by Jackson-Grusby et al., (2001) are conditionally null for *Dnmt1*$^{-/-}$ whereas the *Dnmt1*$^{-/-}$ MEFs used in this study have been suggested to be severe hypomorphs rather than *Dnmt1* null (Li et al., 1992). In either cell line DNA methylation levels of *LINE1* have not been assessed. It is possible that the hypomorphic *Dnmt1* mutant MEFs are able to maintain DNA methylation of *LINE1* to some extent and therewith transcriptional repression.

The reason for *MMERVK10C* not showing an increase in transcription in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs could be that other or additional mechanisms facilitate silencing of this element in fibroblasts or that *MMERVK10C* requires specific transcription factors which might be present germ cells but not in somatic cells. However, data obtained by Michael Ling show no statistically significant increase of *MMERVK10C* transcript in *Dnmt3L*$^{-/-}$ testes (see 3.3.1). In comparison *IAP* elements are upregulated about 6-fold by affymetrix microarray in *Dnmt3L*$^{-/-}$ testes suggesting that *MMERVK10C* behaves different from *IAP* in *Dnmt3L* mutant testes (Hata et al., 2006). To clearly test if DNA methylation is required for transcriptional repression of *MMERVK10C* in this tissue bisulfite sequencing of the *MMERVK10C* locus and comparison with DNA methylation of *IAP*

and *LINE1* promoter regions, which are mis-expressed in the absence of *Dnmt3L*, would be required. In contrast to *Dnmt3L*$^{+/-}$ and *Dnmt3L* $^{-/-}$ testes and *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ fibroblasts, NIH3T3 fibroblasts showed mis-expression of *MMERVK10C* by qRT-PCR analysis upon Aza treatment compared to untreated controls (figure 3.6).Why *MMERVK10C* loci show demethylation in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ fibroblasts but not Aza treated NIH3T3 cells is unclear as in both cases *IAP* and *Tex19.1* transcripts, which have been reported to show severe upregulation in response to demethylation at their promoters, are strongly mis-expressed (figure 3.6) (Walsh and Bestor, 1999; Gaudet et al., 2004, Hackett et al., submitted). It is possible that only a subset of cells has undergone DNA demethylation upon Aza treatment at the time of harvesting. The expression changes observed by qPCR could result from this subset of cells. The discrepancy between the expression analysis and bisulfite sequencing might be caused by bisulfite sequencing not being sensitive enough to detected DNA demethylation in only a subset of cells. Another explanation might be that *MMERVK10C* might be very efficiently targeted by *Dnmt1* or the *de novo* methyltransferases and hence Aza treatment over a limited period of time might not be sufficient to induce demethylation at *MMERVK10C* loci. Bisulfite sequencing of the *Tex19.1* and *IAP* promoters should be performed and compared to the methylation status of the *MMERVK10C* promoter in Aza treated cells. No or little demethylation of *Tex19.1* and *IAP* loci would suggest that the expression changes observed in figure 3.6 result from a subset of cells in the population and that this subset of cells is not sufficient to show DNA demethylation by bisulfite sequencing of the whole cell population. Alternatively, analysis of the DNA methylation status might not correspond to the same copies that are examined on the RNA level due to the large number of copies in the mouse genome. Taken together the data I presented here show that *MMERVK10C* behaves different from *IAP* elements, which also belong to the ERVK family, in fibroblasts in response to DNA demethylation. When I tested if *Tex19.1* expression might modulate the levels of *MMERVK10C* transcript in *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ cells I found that *MMERVK10C* was not differentially expressed between *Tex19.1*$^{-/-}$ and control Aza treated pMEFs (figure 3.6). This suggests that the lack of *MMERVK10C* mis-expression in *p53*$^{-/-}$*Dnmt1*$^{-/-}$ cells is not mediated by the upregulation

of *Tex19.1*. In conclusion *MMERVK10C* does not rely as heavily on DNA methylation for transcriptional silencing in somatic cells and possibly, based on the *Dnmt3L* data from Michael Ling, in germ cells as *IAP* elements. As mentioned above Ian Adams did not observe hypomethylation of *MMERVK10C* LTR/5'UTR in *Tex19.1*$^{-/-}$ testes. This and the here presented data suggest that *MMERVK10C* mis-expression during spermatogenesis in the absence of *Tex19.1* is caused by other mechanisms than transcriptional upregulation through hypomethylation of this element.

3.3.3 Transcriptional regulation of *MMERVK10C* by histone modifications

As mentioned above Aza treatment of NIH3T3 fibroblasts induced expression of *MMERVK10C* transcript which was surprising given the fact that *MMERVK10C* did not appear to be upregulated *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ fibroblasts. One explanation could be that Aza facilitates, in addition to DNA demethylation, changes in histone modifications. It is known that treatment with Aza results in a decrease of the repressive histone mark H3K9me^2 and it is possible that Aza treatment affects further histone modifications like H3K9me^3 (Wozniak et al., 2007). This would be in accordance with the observations made in figures 3.5 and 3.6 which showed upregulation of *MMERVK10C* RNA in Aza treated fibroblasts but not *p53*$^{-/-}$ *Dnmt1*$^{-/-}$ MEFs. It is known from computational methods for analysis of repetitive elements from short-read sequencing data that retrotransposons belonging to the ERVK family, like MMERVK10C and IAPs, generally appear within clusters in the genome enriched for H3K9me3 (Day et al., 2010). In accordance with this it has been previously demonstrated that loss of H3K9me^3 strongly correlates with induction of *MMERVK10C* expression in ES cells null for the methyltransferase *Setdb1* (Karimi et al., 2011). In contrast, *Dnmt* TKO ES cell lines only show a slight induction of *MMERVK10C* transcript and simultaneous knockdown of *Setdb1* and *Dnmt1* in ES cells by RNA interference did not increase the level of *MMERVK10C* expression over that observed upon loss of *Setdb1* alone. In contrast to *MMERVK10C* young *IAP*

elements show an accumulative effect of de-repression upon loss of H3K9me^3 and DNA methylation in pluripotent cells (Rowe el al., 2010; Karimi et al., 2011). This strengthens the hypothesis that *MMERVK10C* is repressed in a different way from *IAP* elements. It appears that H3K9me^3 is one of the major repressors of *MMERVK10C* activity in ES cells. In ES cells, Setdb1 is recruited to target loci via Kap1, and as mentioned before mutations in *Kap1* result in global loss of H3K9me^3 and retrotransposon mis-expression, including MMERVK10C, in ES cells (Rowe et al., 2010). H3K9me^3 is also a major contributor to silencing of *IAP* elements in ES cells and early embryos (Rowe et al., 2010, Karimi et al., 2011). Upregulation of *IAP* RNA in *Dnmt* TKO ES cells is a modest ~4-8 fold compared to the mis-expression observed in hypomethylated somatic cells (Hutnick et al., 2010). Taken together, those data suggest that in contrast to *IAPs*, where silencing by H3K9me^3 in ES cells is replaced by DNA methylation as the major repressor in somatic cells, *MMERVK10C* repression does not appear to undergo the same switch upon ES cell differentiation. It is unclear if H3K9me^3 or other mechanisms facilitate silencing of *MMERVK10C* in somatic cells. In addition to DNA methylation and H3K9me3, histone deacetylation has been implicated in transcriptional silencing of certain retrotransposons in human EC cells and mouse somatic cells (Brunmeir et al., 2010; Garcia-Perez et al., 2010). However, histone deacetylation, at least in cases where it depends on HDACs that can be inhibited by TSA, can be excluded to be a major contributor to *MMERVK10C* transcriptional silencing in fibroblasts (figure 3.7). As H3K9me3 seems to be a major determinant of *MMERVK10C* transcriptional repression in pluripotent cells, we tested if *MMERVK10C* might be also upregulated in *Kap1*$^{-/-}$ pMEFs. Figure 3.7 shows that this is not the case. This suggests that mechanisms other that H3K9me^3 silence *MMERVK10C* in somatic cells or that histone methyltransferases are recruited to *MMERVK10C* loci in a Kap1 independent manner in somatic cells. Reagents like somatic cell lines with mutations in genes known to facilitate H3K9me^3 in differentiated cells or small molecule inhibitors of histone methyltransferases would help to understand if H3K9me3 is required to silence *MMERVK10C* in somatic cells. It would be interesting to perform chromatin-immuno-precipitation (ChIP), which can be used to investigate the interaction between proteins and DNA, on *Tex19.1*$^{+/+}$ and

Tex19.1$^{-/-}$ testes for H3K9me^3 at the *MMERVK10C* locus in order to establish if this mark is impaired in the mutants. However, microarray data from *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/+}$ testes showed that only 0.05% (10 probes) out of 19,089 probes were significantly upregulated by at least 2-fold in *Tex19.1*$^{-/-}$ testes and that six of these probes belonged to *MMERVK10C* (Reichmann et al., submitted). Based on those data it seems unlikely that a global loss of this repressive histone mark has occurred in *Tex19.1*$^{-/-}$ mutants. Similarly, it also seems implausible that *Tex19.1* should affect H3K9me^3 solely at the *MMERVK10C* locus. Taken together, despite H3K9me^3 being a possible repressor of *MMERVK10C* transcription, it appears likely that *Tex19.1* promotes silencing of this element by mechanisms other than H3K9me^3.

3.3.4 Post-transcriptional regulation of retrotransposons by *Tex19.1*

Instead of affecting *MMERVK10C* transcription, *Tex19.1* could also influence any of the other stages of the retrotransposon life cycle. Since Tex19.1 is a cytoplasmic protein any direct effects on *MMERVK10C and* other retrotransposons might be post-transcriptional and acting on the RNA or protein rather than the DNA level. Western blot analysis for *IAP*, *IAPE* and *LINE1* protein levels showed that LINE1 ORF1p but not IAP or IAPE protein levels are elevated in 16dpp *Tex19.1*$^{-/-}$ testes compared to controls (figure 3.8). This is not caused by an increase in *LINE1* RNA (figure 3.8). The 3-fold upregulation identified by western blot is similar to the upregulation identified by IF studies on sections from *Tex19.1*$^{-/-}$ and control testes were an about 2-fold upregulation could be observed (figure 3.9). IF analysis showed that LINE1 ORF1p can be found in the pachytene spermatocytes which is the cell type where defects in meiotic progression are observed for the first time in *Tex19.1*$^{-/-}$ mutants. Taken together those data suggest a loss of translational repression or an increase of LINE1 ORF1 protein stability in *Tex19.1*$^{-/-}$ male germ cells. Western blot and qPCR analysis for IAP, IAPE and LINE1 protein and RNA expression in ES cells showed that RNA expression appears to be variable between cell lines for all three retrotransposons and no *Tex19.1* dependent change of

IAP and IAPE protein between knockout and control cell lines could be observed (figure 3.10). As mentioned under 3.2.3.2 this discrepancy between the testes and ES cell data sets could be caused by the fact that ES cells might generally show high variation of LINE ORF1p expression levels between cell lines which would make it difficult to pick up small changes in protein abundance. In order to examine if *LINE1* is a target for translational repression by Tex19.1 I transfected reporter plasmids for 5'UTR mediated translation of *LINE1* RNA into *Tex19.1*$^{-/-}$ and *Tex19.1* $^{+/+}$ ES cells. This showed no difference between mutant and control cell lines (figure 3.11). This suggests that Tex19.1 does either not regulate the translation of *LINE1* via its 5'UTR or that Tex19.1 is generally not a translational repressor of *LINE1* in ES cells. However, given the large amount of diverse *LINE1* elements, in particular three active subclasses, in the mouse genome it is also possible that Tex19.1 acts on some but not other subclasses and that the sequence of the *LINE1*$_{spa}$ element, belonging to the Tf-type of *LINE1s*, used in this study is not a target for Tex19.1. Overall further work is required to establish if *Tex19.1* might have similar roles in testes and embryonic stem cells. In both, wild type testes and ES cells, Tex19.1 does not associate with polysomes (figure 3.14). Furthermore the data presented in figure 3.12 suggest that Tex19.1 is not associated with polyadenylated mRNA in wild type testes and neither *LINE1* nor *MMERVK10C* RNA shift into the more translated part of the sucrose gradient in *Tex19.1*$^{-/-}$ compared to control testes under the conditions used during this project (figure 3.15). Taken together the here presented data suggest that Tex19.1 is not a translational regulator of *LINE1* or *MMERVK10C*. The discrepancy between the result presented in figure 3.15 compared to the shift of *MMERVK10C* into the more translated part of the sucrose gradient in the absence of *Tex19.1* observed by Ian Adams (figure 3.1) might be based on the fact that the qPCR primers amplify a different subset of *MMERVK10C* transcripts compared to the northern blot probe. Another possible explanation might be that, as can be seen from figure 3.1, *MMEVRK10C* in the northern blot analysis can predominantly be found in fraction 7, however, the polysomes localise to fractions 5 to 10. If *MMERVK10C* RNA would be shifting in response to translation a more even distribution might be expected. The fact that *MMERVK10C* is found in the heavier compartment of the sucrose gradient

by northern blot could be caused by an increase of viral particle formation in the absence of *Tex19.1* as viral particles are relatively large protein assemblies (Hammarstedt et al., 2000). For the qPCR analysis individual mRNP, monosome and polysome fractions were pooled together based on the recorded profile. If an increase in viral particles should be responsible for the shift, which possibly results in an enrichment of retrotransposon RNA predominantly in a certain fraction of the heavier part of the sucrose gradient, in *Tex19.1* mutants, pooling of fractions might dilute this effect. However, the fact that *MMERVK10C gag*, *pol* and *env* transcripts are all upregulated to different extents as shown in figure 3.4 represents a possible caveat with this theory. In order to test if either of the two theories might be true individual fractions of the gradients should simultaneously be analysed by qPCR and northern blot and then compared.

3.3.5 Tex19.1 interacts with Ubr2 and may modulate N-end rule mediated protein turn over

In order to understand by which mechanism Tex19.1 might represses retrotransposon expression we wanted to examine if Tex19.1 functions in a complex and with which proteins it might interact. Figure 3.16A shows that the majority of Tex19.1 is likely bound in a complex as it migrates at larger sizes than its predicted molecular weight on a sucrose size gradient. Interestingly, we were able to confirm Ubr2 being a binding partner for Tex19.1 in ES cells (figure 3.16B). Interaction between Tex19.1 and Ubr2 in male germ cells and Tex19.1 and UBR2 in HeLa cells has been previously reported (Yang et al., 2010). Figure 3.17 shows that Tex19.1 enhances Ubr2 activity in a gain of function experiment but not a loss of function experiment. In addition to Ubr2 there are believed to be at least 3 other E3 proteins that recognize N-degrons (Tasaki et al., 2005). Therefore one possible explanation of the discrepancy between the loss of function experiment in the ES cells and the gain of function experiment in the HeLa cells could be that other Ubr proteins in the *Tex19.1*$^{-/-}$ ES cell lines target N-end rule GFP substrates efficiently enough that no difference between wild type and controls can be

observed. In conclusion the data I present here demonstrate that the interaction between Tex19.1 and Ubr2 is conserved in ES cells which suggests that Tex19.1 might fulfil similar roles in germ cells and pluripotent cells. Based on the data presented in figure 3.17D I speculate that Tex19.1 might promote Ubr2 function. However, further work will be required to establish if this indeed the case and importantly which proteins represent possible targets for the Tex19.1 and Ubr2 complex.

3.3.6 Functional role of the Tex19.1 protein in retrotransposon repression

It was previously known that *MMERVK10C* transcript is upregulated in *Tex19.1*$^{-/-}$ testes but it was not clear if retrotransposon upregulation might be more widespread (Öllinger et al., 2008). Here I have advanced our understanding of this and shown that retrotransposon mis-expression in *Tex19.1*$^{-/-}$ testes is fairly restricted to *MMERVK10C*. Furthermore, I have provided evidence that DNA methylation appears to play a minor role in silencing *MMERVK10C* at the DNA level. How relevant this might be for the RNA up-regulation of *MMERVK10C* observed in *Tex19.1*$^{-/-}$ male germ cells remains to be seen as I was able to provide evidence that Tex19.1 might not act at the DNA, RNA or translational level but at the level of protein stability. LINE1 ORF1p appears to be upregulated in *Tex19.1*$^{-/-}$ testes and based on the interaction between Tex19.1 and Ubr2 it is possible that in the absence of Tex19.1 LINE ORF1p is stabilised due to a loss of N-end rule mediated proteolysis. However, it is currently not clear if LINE ORF1p represents indeed a substrate for Ubr2. It is possible that cleavage of a few amino acids from the LINE1 ORF1p N-terminus creates N-degrons and makes LINE1 ORF1p a target for Ubr2. Alternatively Tex19.1 might act as a scaffold/adaptor to help Ubr2 ubiquitinate LINE1 ORF1p. Currently this is merely speculation and further analysis of retrotransposon proteins, including MMERVK10C, will be required to understand the role of the interaction between Tex19.1 and Ubr2 and the *Tex19.1*$^{-/-}$ male phenotype.

Chapter 4: Characterisation of the role of *Tex19.1* during mouse embryonic development

4.1 Introduction

Tex19.1 expression has been reported to be restricted to germ cells, pluripotent cells and the placenta (Childs, 2006; Kuntz et al., 2008; Öllinger et al., 2008). In the germline *Tex19.1* RNA expression has been reported in migrating PGCs isolated from the hind-gut region of 9.5dpc embryos by qRT-PCR (Hackett et al., submitted). The same study demonstrated *Tex19.1* expression at 10.5dpc and 13.5dpc in the gonads of both sexes. This is in accordance with data where RNA expression of *Tex19.1* from 13.5dpc through to adulthood in the gonads of both sexes has been reported by RT-PCR (Childs, 2006; Kuntz at al., 2008). During embryonic development *Tex19.1* RNA has been shown to be expressed at the 1, 2, 8 and 16 cell and blastocyst stage (Kuntz et al., 2008). RT-PCR showed that *Tex19.1* is also expressed in embryonic germ cells (EG) and ES cells and upon *in vitro* differentiation *Tex19.1* becomes down-regulated in ES cells (Childs, 2006; Kuntz et al., 2008). Furthermore, Tex19.1 protein expression has been demonstrated in fetal germ cells at 14.5dpc, and in male germ cells in the adult gonads as well as in ES cells (Öllinger et al., 2008). This expression pattern is similar to *Oct4*, as described in chapter 1**.**

In order to analyze the physiological function of *Tex19.1*, Rupert Öllinger generated knockout mice. Generation and analysis of mice that have disruption in one or more genes is a powerful reverse genetic approach for studying the role of genes whose function has not been determined. In the case of the *Tex19.1* gene, mutants were generated by replacing the *Tex19.1* open reading frame with a neomycin selection cassette by homologous recombination in E14 embryonic stem cells (figure 4.1) (Öllinger et al., 2008).

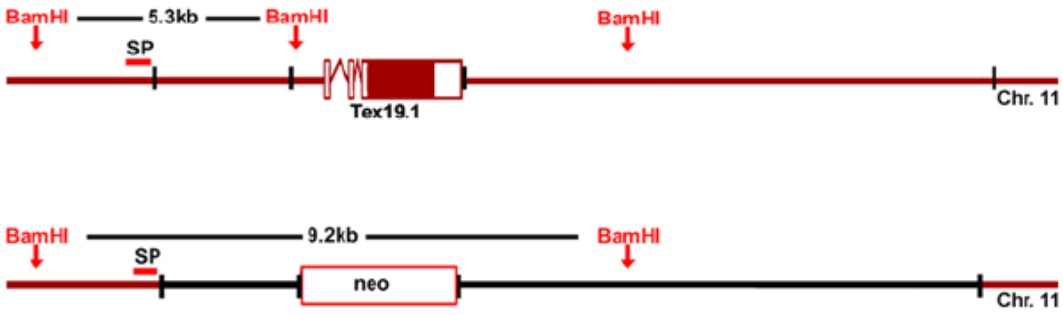


Figure 4.1: Schematic representation of the *Tex19.1* locus. In order to generate *Tex19.1* knockout mice the *Tex19.1* gene in the wild type allele is replaced with a neomycin cassette. SP indicates the location used for the Southern blot probe and BamHI sites and mark the length of restriction fragments for each allele. Taken from Öllinger et al., (2008).

By crossing animals heterozygous for the knockout allele Rupert Öllinger identified a reduction in the number of *Tex19.1*$^{-/-}$ animals born on a mixed genetic background (129/Ola x CD1). The second aim of my thesis was to characterise the embryonic lethality of *Tex19.1*$^{-/-}$ mice in more detail.

During this project I repeated the breeding analysis with mice where the *Tex19.1* mutation had been back crossed three times into a C57BL/6 background. The reasons for this were two-fold. The 129/Ola x CD1 background used in the original study could not be easily maintained since genetic and therewith phenotypic variability would increase in subsequent generations due to the hybrid background. Secondly, the backcross into a C57BL/6 background reduced genetic variation between animals. In order to understand at which stages of embryonic development absence of *Tex19.1* might affect viability, I characterised the expression pattern of *Tex19.1* in the embryo in more detail. As mentioned above *Tex19.1* is expressed in the pluripotent cells of the early pre-implantation embryo. In order to investigate if *Tex19.1*$^{-/-}$ knockouts die due to defects in those cells and in order to establish an *in vitro* model for the characterization of *Tex19.1* function we analyzed *Tex19.1*$^{-/-}$ knockout ES cells. Mouse ES cells are, as mentioned before, derived from the ICM of the blastocyst and can therefore be considered as a model for the pluripotent cells of the pre-implantation embryo. Ian Adams and Rupert Öllinger previously derived ES cells from *Tex19.1*$^{+/-}$ heterozygous crosses. They found

that among 20 generated ES cell lines there was no significant deviation from the expected Mendelian ratio (4 *Tex19.1*$^{-/-}$ ES cells, 6 *Tex19*$^{+/+}$ and 10 *Tex19.1*$^{+/-}$). In order to examine if the reduction of *Tex19.1*$^{-/-}$ homozygotes is caused by problems in the pluripotent cells of the pre-implantation embryo I characterized if *Tex19.1*$^{-/-}$ ES cells can differentiate and if they show upregulation of retrotransposons. Furthermore, I analysed the karyotype of *Tex19.1*$^{-/-}$ ES and control cell lines in order to understand if an increase in aneuploidy in those cells could be causing the phenotype.

4.2 Results

4.2.1 *Tex19.1*$^{-/-}$ homozygotes are born at a reduced frequency

Analysis of heterozygous crosses in a C57BL/6 background showed a 32% reduction of *Tex19.1*$^{-/-}$ homozygotes born. I found that of 785 pups born from heterozygous crosses, 213 were *Tex19.1*$^{+/+}$, 442 *Tex19.1*$^{+/-}$ and 130 *Tex19.1*$^{-/-}$. This is a significant deviation from the expected 1:2:1 ratio (χ^2-test $p < 0.0005$) and suggests that *Tex19.1* plays a role during embryonic development (summarized in table 4.1).

Genotype	Number (and %) of pups observed	Number (and %) of pups expected
Tex19.1$^{+/+}$	213 (27)	193 (25)
Tex19.1$^{+/-}$	442 (56)	386 (50)
Tex19.1$^{-/-}$	130 (17)	193 (25)

Table 4.1: Observed and expected distribution of *Tex19.1*$^{-/-}$ pups.

4.2.2 Reduction of *Tex19.1*$^{-/-}$ homozygotes is exacerbated in subsequent litters and shows strong male/female sex bias

Interestingly, further analysis of the breeding data revealed that the reduction of *Tex19.1*$^{-/-}$ pups is exacerbated in subsequent litters compared to first litters. Table 4.2 shows that pups born from first litters, meaning pups born from female mice that have not given birth before, show a 25% reduction of *Tex19.1*$^{-/-}$ animals (χ^2-test $p<0.05$). In contrast pups born as subsequent litters, which refers to pups born from female mice that have at least given birth once before, and importantly from matings where the mating pair has been constantly kept together, the reduction of *Tex19.1*$^{-/-}$ mutants is approximately 74% and ($p<0.0005$) (table 4.3).

Furthermore, the reduction of homozygote *Tex19.1*$^{-/-}$ animals in subsequent litters is associated with a strong male/female sex bias with females showing an 88% reduction and males a 44% reduction from the expected Mendelian ratio (χ^2-test $p<0.0005$ and $p<0.05$, respectively). This contrasts with pups born from first litters, whereby there is a 27% reduction in homozygote females and a 23% reduction in *Tex19.1*$^{-/-}$ males. This effect was also apparent on the mixed genetic background as observed by Rupert Öllinger (data not shown). Taken together, this data suggests that absence of *Tex19.1* affects embryonic development of both sexes and that this effect is exacerbated in pups born as subsequent litters. Furthermore, the increase in embryonic lethality of *Tex19.1*$^{-/-}$ pups in subsequent litters affects females even stronger than males.

Genotype	Total number (and %) of pups observed	Number (and %) of females observed	Number (and %) of males observed	Number (and %) of males or females expected
Tex19.1$^{+/+}$	111 (25)	52 (12)	59 (13)	55.5 (12.5)
Tex19.1$^{+/-}$	250 (56)	120 (27)	130 (29)	111 (25)
Tex19.1$^{-/-}$	83 (19)	40 (9)	43 (10)	55.5 (12.5)

Table 4.2: Observed and expected distribution of *Tex19.1*$^{-/-}$ pups in first litter.

Genotype	Total number (and %) of pups observed	Number (and %) of females observed	Number (and %) of males observed	Number (and %) of males or females expected
Tex19.1$^{+/+}$	60 (30)	25 (13)	35 (17)	25.25 (12.5)
Tex19.1$^{+/-}$	125 (62)	73 (36)	52 (26)	50.5 (25)
Tex19.1$^{-/-}$	17 (8)	3 (1)	14 (7)	25.25 (12.5)

Table 4.3: Observed and expected distribution of *Tex19.1*$^{-/-}$ pups in subsequent litters.

Female mice are receptive to mate shortly after the female has given birth during the post-partum estrus. If a mating pair is consonantly kept together, as described above for pups born as subsequent litters, mating will usually occur during the post-partum estrus (Gilbert, 1984). This means that the female will conceive another pregnancy while suckling a litter. Therefore the data presented in table 4.2 and 4.3 implied to us that the exacerbated reduction of *Tex19.1*$^{-/-}$ pups born from subsequent litters is caused by lactating mothers. I wanted to confirm that viability of *Tex19.1*$^{-/-}$ embryos was indeed compromised by lactation of the mother and not caused by some other defect that might occur in subsequent pregnancies in *Tex19.1*$^{+/-}$ females. In order to do so, I analysed pups born from subsequent litters where the mating pair had been separated as soon as the female was plugged. This procedure ensured that the female would not be suckling a litter when becoming pregnant again. The numbers of homozygote pups born from subsequent litters of non-lactating females were similar to those born from first litters (χ^2-test $p>0.2$) (tables 4.2, 4.3 and 4.4). This contrasts significantly with the number of pups born from lactating mothers and therefore suggests exacerbated lethality of *Tex19.1*$^{-/-}$ embryos occurs indeed in response to lactation (χ^2-test $p< 0.0005$) (tables 4.2, 4.3 and 4.4).

Genotype	Total number (and %) of pups observed	Number (and %) of females observed	Number (and %) of males observed	Number (and %) of males/females expected
Tex19.1$^{+/+}$	46 (30)	23 (15)	23 (15)	19.25 (12.5)
Tex19.1$^{+/-}$	76 (49)	34 (22)	42 (27)	38.5 (25)
Tex19.1$^{-/-}$	32 (21)	11 (7)	21 (14)	19.25 (12.5)

Table 4.4: Observed and expected distribution of *Tex19.1*$^{-/-}$ pups in subsequent litters separated.

As described in chapter 1, in lactating mothers, embryos in the next litter usually enter a developmental delay at the blastocyst stage, known as diapause in order to prevent two

litters being born too close to each other (Renfree and Shaw, 2000). In order to test if *Tex19.1* might be required during embryonic diapause, we chemically induced developmental arrest with tamoxifen and depo-provera as described previously (Hunter and Evans, 1999). In each case 15 females were superovulated, mated and at 2.5dpc injected with tamoxifen and depo-provera to induce diapause. At 6.5 dpc the arrested blastocysts were isolated and transferred to the uterus of 2.5 dpc pseudopregnant recipient females (the isolation of blastocysts and embryonic transfers were carried out by Emma Murdoch). In the two experiments, only 63 and 70 blastocysts were isolated and transferred and from those, 23 and 25 pups were born and genotyped, respectively. Surprisingly, independent of genotype, I noticed a 54% reduction in the number of females born from chemically diapaused blastocysts (χ^2-test $p<0.0005$). Since the major reduction of *Tex19.1*$^{-/-}$ pups occurs in females, this experiment is not appropriate to address the question of whether or not *Tex19.1* plays a role during embryonic diapause. This experiment would need to be repeated at least twice and therefore the number of mice required to obtain a sufficient number of chemically arrested male and female pups born appears unreasonable. Accordingly, we did not pursue this experiment further. Instead, we analysed lactation-induced diapaused embryos in more detail.

Genotype	Total number (and %) of pups observed	Number (and %) of females observed	Number (and %) of males observed	Number (and %) of males/females expected
Tex19.1$^{+/+}$	11 (23)	2 (4)	9 (19)	6 (12.5)
Tex19.1$^{+/-}$	30 (62.5)	8 (16)	22 (46)	12 (25)
Tex19.1$^{-/-}$	7 (14.5)	1 (2)	6 (13)	6 (12.5)

Table 4.5: Observed and expected distribution of *Tex19.1*$^{-/-}$ pups born from chemically diapaused embryos.

To investigate at which stages mutation in *Tex19.1* affects viability of embryos in lactating mothers, diapaused embryos were isolated at the equivalent of 13.5 dpc from

nursing females, genotyped and sexed by PCR. All embryos isolated had undergone diapause which varied between 3 to 5 days from litter to litter. However, difficulties arose through the fact that it appeared that C57BL/6 females may have problems to maintain embryonic diapause in general as 40% of investigated females were not pregnant at the desired stage but showed degenerated implantation sites. However, it is not clear if this is indeed true for C57BL/6 females as there have not been any studies that have examined the ability of different mouse strains to maintain diapause to date. No deformed or developmentally retarded embryos were observed among the isolated embryos, however, as table 4.6 shows, the reduction of a large amount of *Tex19.1*$^{-/-}$ female embryos has occurred by 8.5-10.5 dpc (χ^2-test $p<0.05$). During the process of isolating the embryos from the deciduas for genotyping, 13 bloody and/or empty deciduas were noticed. This suggests that a large proportion of *Tex19.1*$^{-/-}$ embryos has died before 10.5dpc. In order to test if *Tex19.1*$^{-/-}$ embryos suffered lethality during embryonic diapause or after implantation, lactation-induced diapaused blastocysts were isolated at the equivalent of 5.5dpc and genotyped. All isolated and genotyped blastocysts looked morphological normal with epiblast, hypoblast and trophectoderm being clearly distinguishable by brightfield microscopy (not shown). Gene expression differences between *Tex19.1*$^{-/-}$, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{+/+}$ blastocysts were not investigated in this study. Table 4.7 shows that distribution of *Tex19.1*$^{+/+}$, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ embryos among lactation-induced diapaused blastocysts is not different from the distribution observed in pups born as first litters (χ2-test $p>0.5$) but significantly different from pups born as subsequent litters (χ2-test $p<0.05$). This suggests that the reduction we observe in *Tex19.1* homozygote embryos does not occur at the blastocyst stage after 1-2 days of diapause but likely around or after implantation.

Genotype	Total number (and %) of embryos observed	Number (and %) of females observed	Number (and %) of males observed	Number (and %) of males or females expected
Tex19.1$^{+/+}$	27 (28)	14 (14)	13 (13.5)	10.875 (12.5)
Tex19.1$^{+/-}$	54 (56)	28 (29)	26 (27)	21.75 (25)
Tex19.1$^{-/-}$	16 (16)	4 (4)	12 (12.5)	10.875 (12.5)

Table 4.6 Observed and expected distribution of embryos isolated from lactating females at 13.5dpc. Developmental stage of embryos appeared to vary from litter to litter. It appears that the embryos analysed here underwent between 3 to 5 days of diapause.

Genotype	Total number (and %)of embryos observed	Number (and %) of embryos expected
Tex19.1$^{+/+}$	5 (17)	7.25 (25)
Tex19.1$^{+/-}$	18 (62)	14.5 (50)
Tex19.1$^{-/-}$	6 (21)	7.25 (25)

Table 4.7 Observed and expected distribution of lactation induced diapaused blastocysts.

4.2.3 *Tex19.1* expression is associated with a pluripotent cell state in embryonic cell lineages and is also present in extraembryonic tissues during development

In order to understand why viability of embryos is affected when Tex19.1 is absent we examined the expression pattern of *Tex19.1* in more detail. It has been previously reported that *Tex19.1* RNA becomes down regulated upon ES cell differentiation (Childs, 2006; Kuntz et al., 2008). To see if *Tex19.1* RNA expression is indeed associated with pluripotency in embryonic stem cells we differentiated E14 ES cells in hanging droplet culture (figure 4.2). RT-PCR showed that *Tex19.1* RNA expression becomes down regulated upon differentiation and is not detectable by day 7 of differentiation. The down-regulation of *Tex19.1* coincides with down-regulation of the pluripotency markers *Oct4* and *Nanog* and up-regulation of markers for the 3 germ

layers (*Gata4* endoderm, *Msx3* ectoderm, *Nkx2.5* mesoderm). This suggests that *Tex19.1* is expressed in pluripotent cells and becomes down-regulated upon differentiation. As mentioned before, it has been shown that *Tex19.1* is expressed in the pluripotent cells of the pre-implantation embryo (Kuntz et al., 2008). However, the expression of *Tex19.1* in the post-implantation embryo has not been reported but there are pluripotent cells present also at early stages of this process. In order to see if *Tex19.1* is expressed after implantation I performed immunohistochemistry (IHC) on sections of early post-implantation embryos. Figure 4.3 shows that anti-Tex19.1 staining can be observed in the early post-implantation embryo. At 6.5dpc anti-Tex19.1 staining is present in the ectoplacental cone, extraembryonic ectoderm and the epiblast. In 7.5 dpc embryos Tex19.1 staining is present in extraembryonic tissues but is downregulated in epiblast-derived tissues in the embryo with some faint staining present in embryonic mesoderm and extraembryonic mesoderm. One day later, Tex19.1 is not detectable in the embryonic tissues at 8.5 dpc. Tex19.1 staining is also not present in the hindgut endoderm where primordial germ cells are located at this stage. At 9.5dpc Tex19.1 staining is mainly detected in the placenta. Concerning the observed reduction of *Tex19.1*$^{-/-}$ homozygotes born from heterozygous crosses, the expression data presented here, together with the published data, suggest that the phenotype could result from problems in the pluripotent cells of the pre-implantation embryo, problems in the pluripotent cells or from defects in the function or development of the extraembryonic cell lineages.

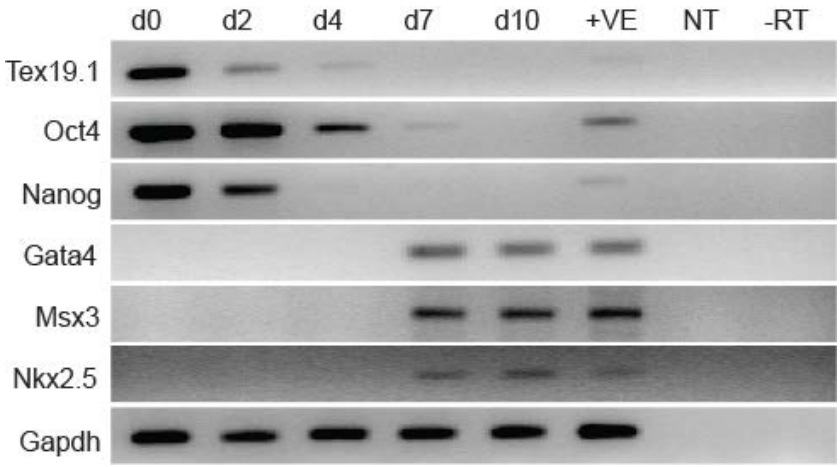


Figure 4.2: *Tex19.1* expression is downregulated upon ESC differentiation. ESCs were differentiated by hanging droplet culture in the absence of LIF. Loss of pluripotency is reflected by loss of *Oct4* and *Nanog* expression. *Tex19.1* expression is not detectable by day 7 of differentiation, which is in accordance with the loss of pluripotency. Expression of *Gata4*, *Msx3* and *Nks2.5* marks differentiation into the three germ layers (d0 is 0 days of differentiation, d2 is second day of differntiation etc., +VE is RT-PCR for whole 9.5dpc embryo, NT is no template control, -RT is minus reverse transcriptase control).

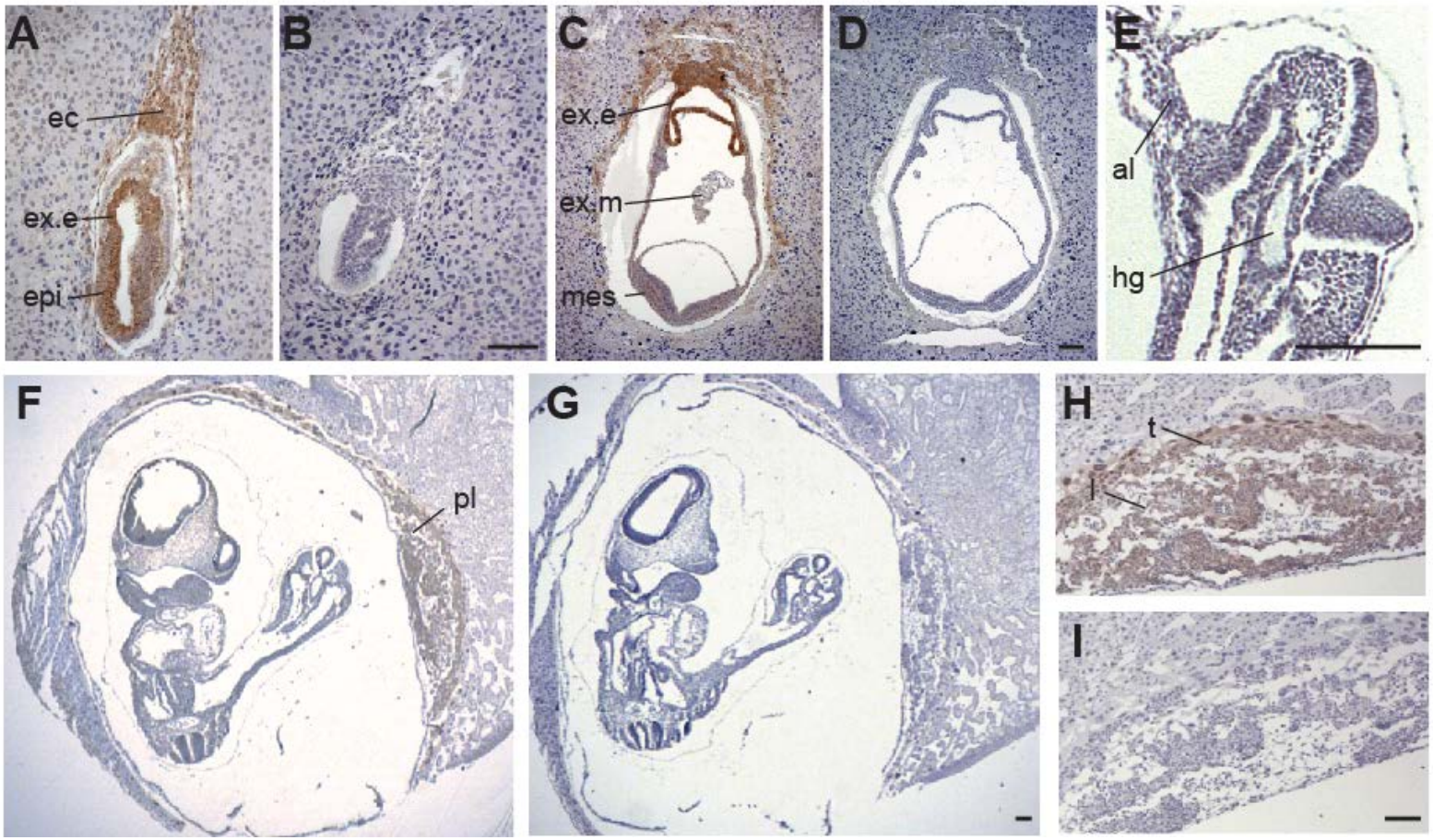


Figure 4.3: Tex19.1 protein expression in post-implantation embryos. (A) Anti-Tex19.1 staining (brown precipitate) in 6.5 dpc embryos is present in the ectoplacental cone (ec), extraembryonic ectoderm (ex.e), and epiblast (epi). (B) 6.5 dpc IgG control. (C) In 7.5 dpc embryos Tex19.1 staining is present in extraembryonic tissues but is downregulated in epiblast-derived tissues in the embryo with some faint staining present in embryonic mesoderm (mes) and extraembryonic mesoderm (ex.m). (D) 7.5 dpc IgG control.

(E) Tex19.1 staining was not detectable in embryonic tissues at 8.5 dpc. Tex19.1 staining was not detected in the hindgut endoderm where primordial germ cells are located at this stage. Panel shows posterior region of 8.5 dpc embryo. hg, hindgut; al, allantois. (F) At 9.5 dpc Tex19.1 staining is detected mainly in the placenta (pl). (G) 9.5 dpc IgG control. (H) Higher magnification of the placenta at 9.5 dpc, Tex19.1 staining is present in the trophoblast (t) and future labyrinthine (l) layers. (I) 9.5 dpc placenta IgG control. Scale bars are 100 microns.

4.2.4 Analysis of *Tex19.1*$^{-/-}$ ES cells

In order to analyse if the reduction of *Tex19.1*$^{-/-}$ pups might be caused by a problem in the pluripotent cells, we used embryonic stem cells as a model system. To test if the embryonic stem cell lines derived by Ian Adams and Rupert Öllinger expressed pluripotency markers western blot analysis was performed on extract from *Tex19.1*$^{+/+}$, $^{+/-}$ and $^{-/-}$ ES cells. Figure 4.4 shows that Oct4 expression levels appear similar between *Tex19.1*$^{+/+}$, *Tex19.1* $^{+/-}$ and *Tex19.1*$^{-/-}$ ES cells suggesting that absence of *Tex19.1* does not affect Oct4 protein levels in ES cells. To test if ES cells have the ability to differentiate in the absence of Tex19.1, *Tex19.1*$^{-/-}$ ES cells were differentiated by hanging droplet culture. Figure 4.5 shows that *Tex19*$^{-/-}$ knockout ES cells down regulate pluripotency markers and upregulate markers of the 3 germ layers upon differentiation suggesting that Tex19.1 is not required for ES cell differentiation in this assay.

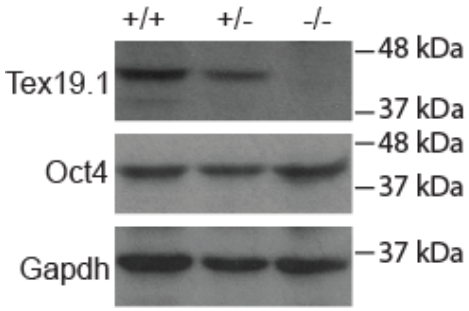


Figure 4.4: Oct4 expression in *Tex19.1*$^{-/-}$ ES cells. For Western blot analysis 20µg of ES cell extract was loaded in each lane. *Tex19.1* $^{+/+, +/-}$ and $^{-/-}$ ES cells were analysed for Tex19.1 and Oct4 protein expression.

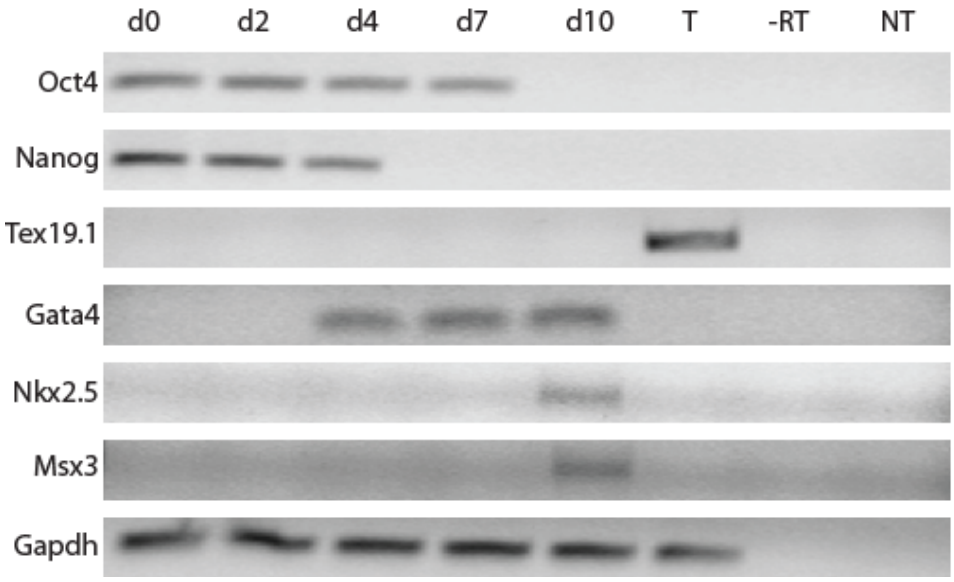


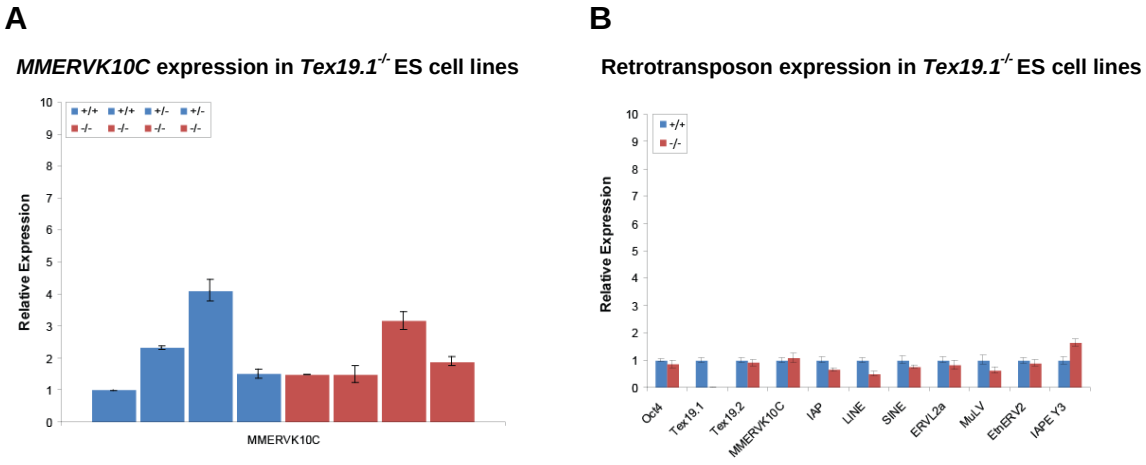
Figure 4.5: *Tex19*$^{-/-}$ knockout ES cells can differentiate into the three germ layers in vitro. ESCs were differentiated by performing hanging droplet culture in the absence of LIF. Loss of pluripotency is indicated by loss of *Oct4* and *Nanog* expression. RT-PCR on testis cDNA was performed as a positive control for Tex19.1. Expression of *Gata4*, *Msx3* and *Nks2.5* marks differentiation into the three germ layers (d0 is 0 days of differentiation, d2 is 2 days of differentiation etc., T is testis RT-PCR, -RT is minus reverse transcriptase control, NT is no template control).

Although we could detect no obvious abnormalities in the derivation, maintenance or differentiation of *Tex19.1*$^{-/-}$ ES cells, I investigated whether loss of Tex19.1might cause de-repression of retrotransposons in ES cells, as MMERVK10 RNA is upregulated in *Tex19.1*$^{-/-}$ male germ cells. According to the fact that pluripotent cells in the early embryo are also a viable target for endogenous retrovirus activity we wanted to know if the loss of *Tex19.1*$^{-/-}$ homozygotes results from an upregulation of retrotransposon in those cells. Retrotransposons in mammals are divided into three major classes (Maksakova et al., 2006) which are:

1. Autonomous long interspersed elements (LINEs)
2. Non-autonomous short interspersed elements (SINEs) and
3. Endogenous retroviruses (ERVs).

We tested *Tex19.1*$^{-/-}$ knockout ES cells by qRT-PCR for upregulation of *MMERVK10C*, a member of each class of retrotransposons and a member of ERV1, ERVK and ERVL

families which represent ERV subfamilies. Furthermore we tested the expression of two retrotransposons closely related to *MMERVK10C*. Figure 4.6A shows that *MMERVK10C* RNA expression does not correlate with presence or absence of *Tex19.1* transcript suggesting that *MMERVK10C* RNA is not mis-expressed in *Tex19.1*$^{-/-}$ ES cells. Similarly, *LINE* and *SINE* retrotransposons show no upregulation in *Tex19.1*$^{-/-}$ES cells (figure 4.6B). *MuLV*, *IAP* and *MERVL2a* LTR retrotransposons representing the ERV1, ERVK and ERVL families of LTR retrotransposons respectively, do not change in expression in *Tex19.1*$^{-/-}$ ES cells either. This appears also to be true for *EtnERV2* and *IAPEY3* elements which are closely related to *MMERVK10C* (figure 4.6B). In addition, I tested a selection of other retrotransposons for expression in *Tex19.1*$^{-/-}$ ES cells compared to controls. Expression of *MysERV*, *ERVB*, *EtERV* and *ETnI* was not increased in *Tex19.1*$^{-/-}$ ES cells (figure 4.6C). In contrast *MusD* appeared to be up-regulated around 7.5-fold in the *Tex19.1*$^{-/-}$ knockout cell line compared to the wild type control. However, analysis of two further control cell lines and two more *Tex19.1*$^{-/-}$ ES cell lines showed that *MusD* expression was up-regulated by only 1.8-fold in one of the *Tex19.1*$^{-/-}$ lines and even downregulated in the other *Tex19.1*$^{-/-}$ line by about 10 fold compared to controls. Furthermore, in one of the controls, the expression of *MusD* appeared downregulated by about 10 fold. This finding suggests that the variation in *MusD* RNA expression between ES cell lines is not linked to the *Tex19.*1 mutation and possibly results from genetic differences between the cell lines.



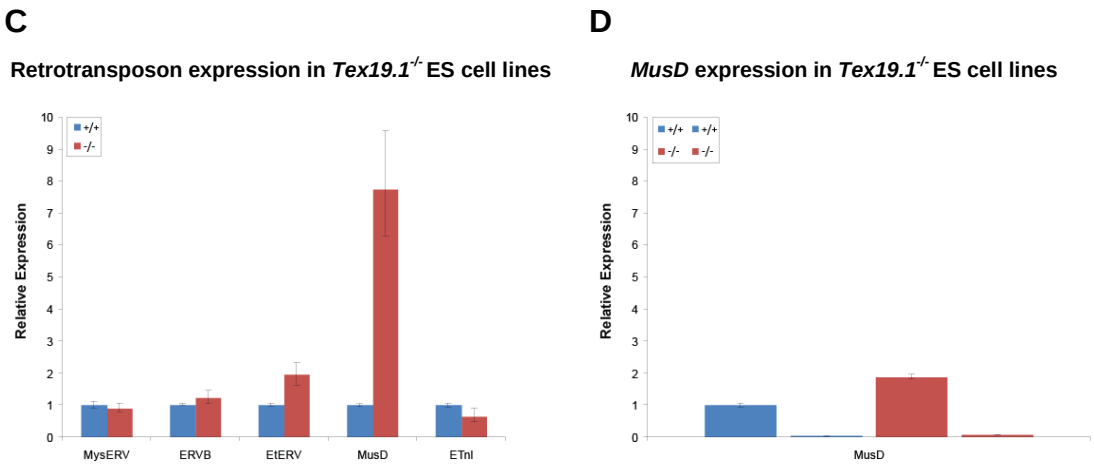


Figure 4.6: Retrotransposon RNA expression in *Tex19.1*$^{-/-}$ ES cells. (A-D) Expression levels for each gene or repetitive element were normalized to β-actin and expressed relative to one wild type control. Each bar represents one biological replicate. Error bars represent standard error among three technical replicates.

In addition to repression of retrotransposon expression, *Tex19.1* has been implicated in promoting faithful chromosome segregation in mouse male and female germ cells (Öllinger et al., 2008; Chapter 5). Therefore I decided to test if *Tex19.1* might play a role in preventing aneuploidy in ES cells. ES cell lines, all of which were between passage number 10-12, were arrested at pro-metaphase and chromosome spreads prepared for 4 *Tex19.1*$^{-/-}$, 2 *Tex19.1*$^{+/-}$ and 2 *Tex19.1*$^{+/+}$ cell lines. Of each cell line, 25 spreads were counted. This revealed high levels of aneuploidy in each of the 4 *Tex19.1*$^{-/-}$ cell lines which were, with a total of 74% percent of spreads being aneuploid, significantly higher compared to the 19% of aneuploidy observed in the *Tex19.1*$^{+/-}$ and Tex19.1$^{+/+}$ controls (χ^2-test $p < 0.0005$).

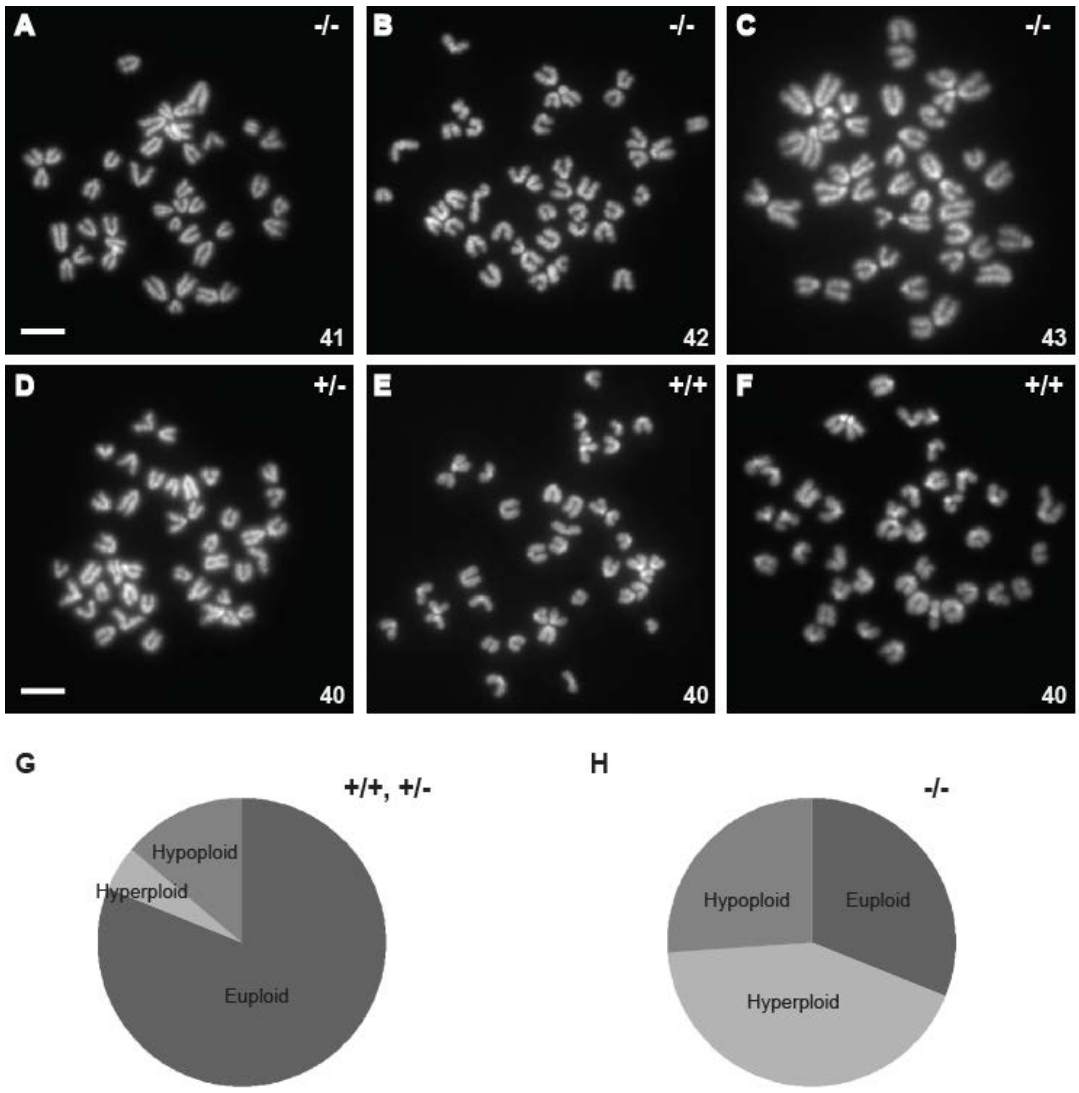


Figure 4.7: Aneuploidy in *Tex19.1*$^{-/-}$ ES cell lines.
(A-C) The majority of cells from 4 different *Tex19.1*$^{-/-}$ knockout ES cell lines show aneuploidy. (A-C) shows chromosome metaphase spreads prepared from *Tex19.1*$^{-/-}$ knockout ES cell lines with 41 (A), 42 (B) and 44 (C) chromosomes. (D-F) Chromosome metaphase spreads prepared from *Tex19.1*$^{+/-}$ heterozygote and *Tex19.1*$^{+/+}$ wild type cell lines are mainly euploid. Scale bars represent 5 µm. (G and H) Among 4 different *Tex19.1*$^{-/-}$ knockout ES cell lines 74% of cells show aneuploidy in contrast to 19% of aneuploidy seen in *Tex19.1*$^{+/-}$ and $^{+/+}$ control ES cell lines (n=100). Only cell lines 35< and >45 chromosomes were taken into account.

4.3 Discussion

In this chapter I set out to gain a better understanding of the role of Tex19.1 during embryonic development. Firstly, I collected breeding data from *Tex19.1*$^{+/-}$ crosses which showed a reduction of *Tex19.1*$^{-/-}$ pups born from the expected ratio which I followed up in detail. Next, I analysed the expression pattern of Tex19.1 during early post-implantation development in order to understand at which stages and cell types mutation in *Tex19.1* might affect embryonic viability, as *Tex19.1* expression patterns had only been reported for pre-implantation development and germ cells so far. Since *Tex19.1* is expressed in pluripotent cells during embryonic development, I tested if *Tex19.1*$^{-/-}$ ES cells could be used as a tool to model *Tex19.1* function during embryonic development. I finished the chapter by providing evidence that absence of *Tex19.1* is associated with aneuploidy in ES cells.

4.3.1 Mutation in *Tex19.1* compromises embryonic viability

Analysis of 785 pups born from *Tex19.1*$^{+/-}$ heterozygote crosses on a C57BL/6 background showed a 33% reduction in the number of *Tex19.1*$^{-/-}$ mutants (table 4.1). This data is similar to the 34% reduction previously reported by our laboratory on a mixed genetic background (29/Ola x CD1). In contrast, Yang et al., (2010) reported an approximately 50% reduction in the number of *Tex19.1*$^{-/-}$ homozygotes on a mixed genetic background born (C57BL/6 x 129/sv). This discrepancy is likely caused by a difference in the experimental set up. The data presented in this thesis suggest that *Tex19.1*$^{-/-}$ animals born from lactating mothers show a stronger reduction of homozygotes than pups born as first litters (tables 4.2 and 4.3). It is possible that variation in the numbers of pups born as first litters and pups born as litters from lactating mothers between the studies accounts for the observed difference. The data presented in this thesis shows that the reduction of *Tex19.1*$^{-/-}$ homozygote increases from 25% in both sexes among pups born as first litters to 44% in males and 88% in females

among pups born from lactating mothers. The fact that lactation is influencing the embryonic lethality is further supported by the observation that pups born as subsequent litters from females that are not lactating, do not show the increase in lethality of *Tex19.1*$^{-/-}$ homozygotes, but are born with a similar distribution as pups from first litters. The reason why embryonic viability is affected in *Tex19.1*$^{-/-}$ mice and why this effect is exacerbated in pups born from lactating mothers, particular females, is not clear. One possibility is that *Tex19.1* plays a role during embryonic diapause (see chapter 1). In order to test if the reduction of *Tex19.1*$^{-/-}$ homozygotes seen in our experiments is caused or exacerbated by blastocysts entering diapause we chemically-induced diapause in embryos from *Tex19.1*$^{+/-}$ matings which, surprisingly, resulted in a 54% reduction of all females born independent of genotype. This suggests that there might be a difference between lactation-induced and chemically-induced diapause, as decreased viability of lactation-induced diapaused female embryos has not been reported. Furthermore, no reduction of *Tex19.1*$^{+/+}$ or *Tex19.1*$^{+/-}$ females was observed in litters that likely had undergone lactation-induced diapause (figure 4.3). It is not clear why chemically-induced diapause would affect viability of female embryos. This observation requires further investigation. As mentioned in chapter 1, little is known about genes required by the embryo itself to undergo diapause, but for example, *gp130*$^{-/-}$ embryos are unable to resume development after diapause as they are not able to maintain the epiblast during developmental arrest (Nichols et al., 2001). In contrast, since *Tex19.1*$^{-/-}$ chemically diapaused embryos were born after embryo transfer it appears that *Tex19.1* is not absolutely required to enter, maintain or exit embryonic diapause. *Tex19.1* expression may or may not influence viability during embryonic diapause. Further analysis showed that a large reduction of *Tex19.1*$^{-/-}$ homozygote females has occurred by 13.5 dpc in 3 to 5 days diapaused embryos isolated from lactating mothers (table 4.6). This reduction, of 63%, is smaller than the reduction observed in table 4.3. This might be caused either by the fact that the number of embryos analysed here is too small, which might bias the data to some extent, or that the lethality of *Tex19.1*$^{-/-}$ homozygotes might not occur at one certain developmental stage but at several stages depending on how severe the individual embryo is affected. Both reasons might also explain the fact that we did not

observe a reduction in the number of *Tex19.1*$^{-/-}$ homozygous males in this experiment. Isolation of lactation-induced diapaused blastocysts showed that reduction of *Tex19.1*$^{-/-}$ embryos does not occur at this stage suggesting that *Tex19.1* homozygote embryos die around or after implantation. It is currently not clear how loss of *Tex19.1* in diapaused embryos would cause lethality after implantation, but the here presented data suggests that expression of *Tex19.1* might be required at a time where a pluripotent stage has to be maintained for a prolonged period. Alternatively, based on the fact that *Tex19.1* is also expressed in the extraembryonic cell lineages, it is possible that not embryonic diapause, but lactation negatively affects embryonic viability of *Tex19.1*$^{-/-}$ mutants (figure 4.3). Survival and growth of the embryo are critically dependent on the extraembryonic cell lineages, in particular the placenta (Cross et al., 2003). The placenta forms the interface between the maternal and fetal circulation and facilitates nutrient uptake, waste elimination, and gas exchange via the mother's blood supply (Watson and Cross, 2005). Genes that affect placental development or function cause fetal growth restriction or embryonic death between 10.5-12.5dpc, often resulting from an impaired nutrient supply (Rashbass et al., 1991; Anson-Cartwright et al., 2000; Chawengsaksophak et al., 2004; Lotz et al., 2004). If absence of *Tex19.1* should cause problems in development or function of the placenta it is possible that this defect affects *Tex19.1*$^{-/-}$ embryos from lactating mothers more severely due to additional restrains on nutrient supply. However, it is unclear why this would affect *Tex19.1*$^{-/-}$ females stronger than *Tex19.1*$^{-/-}$ males. Careful analysis of placental gene expression and histology of *Tex19.1*$^{-/-}$ placenta will be required to understand if defects in the extraembryonic cell lineages might cause the observed reduction of *Tex19.1*$^{-/-}$ embryos.

4.3.2 *Tex19.1* is expressed in pluripotent cells and extraembryonic cell lineages

Figure 4.2 shows that *Tex19.1* RNA expression becomes down-regulated upon ES cell differentiation which suggests that the reduction of *Tex19.1*$^{-/-}$ homozygotes could be caused by a problem in the pluripotent cells during pre-implantation development, as *Tex19.1* RNA expression has been reported to be present from the 1 cell through to the blastocyst stage in pre-implantation embryos (Kuntz et al., 2008). Figure 4.3 also shows that also in the post-implantation embryo, anti-Tex19.1 staining is associated with a pluripotent cell state in the embryo proper as Tex19.1 staining is strongly present at 6.5 dpc in the pluripotent epiblast cells, but is downregulated in epiblast-derived tissues in the embryo proper with some faint staining present in embryonic mesoderm and extraembryonic mesoderm. From 8.5 dpc onwards, anti-Tex19.1 staining is not detectable in the embryonic tissues. The anti-Tex19.1 staining by IHC follows the behaviour of *Tex19.1* RNA expression presented in a study by Hackett et al., (submitted) which shows that *Tex19.1* RNA is expressed in the epiblast at 6.5dpc and becomes down-regulated during gastrulation. *Tex19.1* RNA can only be weakly detected in late 7.5dpc embryos (Hackett et al., submitted). Interestingly, anti-Tex19.1 staining in the post-implantation embryo was not restricted to the embryo proper as staining was also present in the ectoplacental cone and extraembryonic ectoderm. In contrast to the embryonic lineages, anti-Tex19.1 staining does not become downregulated in extraembryonic tissues at 7.5dpc and can be detected at 9.5dpc in the tissue that will give rise to the placenta (figure 4.3). This is in accordance with RT-PCR data that identified *Tex19.1* RNA expression in the placenta (Kuntz et al., 2008). The expression pattern of *Tex19.1* seems similar to the expression pattern of the pluripotency markers *Oct4* and *Nanog*. However, expression of *Oct4* and *Nanog* is more restricted than *Tex19.1* as *Oct4* and *Nanog* are exclusively expressed in pluripotent early embryo cells, the germ cell lineage and pluripotent stem cells but are not expressed in the extraembryonic cell lineages (Schöler et al., 1990; Yeom et al., 1996; Niwa et al., 2000; Chambers et al., 2003; Mitsui et al., 2003). It appears that expression of Tex19.1 is

associated with a pluripotent or germ cell state in embryonic cell lineages but, based on the fact that *Tex19.1*$^{-/-}$ homozygotes are born, albeit at reduced numbers, it is unlikely that *Tex19.1* is absolutely required for pluripotency. This is further supported by the fact that *Tex19.1*$^{-/-}$ ES cells can be derived and maintained at Mendelian numbers (see chapter 4.1). In order to confirm that *Oct4* expression was unaffected in *Tex19.1*$^{-/-}$ mutants I performed western blot analysis. Figure 4.4 shows that *Tex19.1*$^{+/+}$, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ ES cells express similar levels of Oct4 protein suggesting that expression of pluripotency markers is not impaired in the absence of *Tex19.1*. Taken together the data presented so far suggest that the reduction in *Tex19.1*$^{-/-}$ homozygotes is not caused by problems in maintaining a pluripotent state during embryonic development but could nevertheless be caused by problems in the pluripotent cells of the pre- or post-implantation embryo or result from defects in the development and/or function of the placenta. Influenced by the fact that *Tex19.1*$^{-/-}$ ES cells were readily available in the lab at the beginning of this project I decided to characterize those in more detail with the hope that they might provide a valuable tool to study the function of *Tex19.1* during embryonic development and hence provide a system that could help to understand the reduction of homozygote embryos that we see in the *Tex19.1* mutants.

4.3.3 Characterization of *Tex19.1*$^{-/-}$ ES cells

As mentioned above, *Tex19.1*$^{-/-}$ ES cells appear capable of self-renewal, however, embryonic lethality could also be caused by defects in the ability to differentiate. For example embryos with mutations in *Dido3* show defects in differentiation of epiblast cells and ES cells also fail to differentiate (Fütterer et al., 2011). Similarly, *Gata4*$^{-/-}$ mutant ES cells show defects in differentiation upon embryoid body formation and LIF withdrawal as shown by continued expression of *Oct4* (Soudais et al., 1995; Capo-Chichi et al., 2005). As can be seen from figure 4.5, *Tex19.1*$^{-/-}$ ES cells are able to down-regulate pluripotency markers and upregulate markers of all 3 germ layers upon differentiation suggesting that *Tex19.1* is not essential for pluripotent cell differentiation. This result might have been expected as viable *Tex19.1*$^{-/-}$ mice suggest that *Tex19.1*$^{-/-}$

pluripotent cells in the early embryo are able to differentiate. Taken together it appears that derivation, maintenance and differentiation of ES cells is not impaired in the absence of *Tex19.1*. As mentioned before, ES cells are also a viable target for endogenous retrovirus activity and *Tex19.1* has been implicated in retrotransposon repression in male germ cells. Figure 4.6A shows that *MMERVK10C* RNA, which is upregulated at the RNA level in male germ cells, is not mis-expressed in *Tex19.1*$^{-/-}$ ES cells. Similarly, the other retrotransposons analysed show no up-regulation in the absence of *Tex19.1* (figure 4B, C and D). Those data suggest that retrotransposons, at least the studied elements, are not up-regulated at the RNA level in *Tex19.1*$^{-/-}$ ES cells. Genome wide profiling of retrotransposon expression in *Tex19.1*$^{-/-}$ and control ES cells would be useful to strengthen this observation. However, as evidence collected during this project suggests, inhibition of retrotransposons by *Tex19.1* might occur at the protein rather than the DNA or RNA level. It is still not clear if *Tex19.1* has a functional role in ES cells or if embryonic lethality is caused by defects in pluripotent epiblast or extraembryonic cells. Furthermore *Tex19.1*$^{-/-}$ ES cells show high levels of aneuploidy which limits the use of transcriptional profiling.

4.3.4 Aneuploidy in *Tex19.1*$^{-/-}$ ES cell lines

Figure 4.7 shows high rates of aneuploidy in the *Tex19.1* mutant ES cell lines compared to controls. Previous reports suggest that 77% of ES cell lines derived on feeders are initially euploid (Robertson et al., 1983; Robertson and Bradley, 1986). In contrast, data presented by Nichols and colleagues showed that 14 out of 15 derived ES cell lines were characterized by different extents of chromosomal abnormalities, with only one line showing an entirely normal karyotype at 4 or 5 passages when ES cells were derived on gelatin (Nichols et al., 1990). If there is actually an advantage for the use of feeders, it is unclear. It has been reported that differences in aneuploidy levels of ES cell lines in different studies might be caused by particular culture conditions and/or mouse strains used to generate the ES cell (Sugawara et al., 2006). In the same study it was shown that 40% of 88 karyotyped ES cell lines were aneuploid (Sugawara et al., 2006). If those

aneuploid lines have been generated from an aneuploid parental clone or were initially euploid and accumulated aneuploidy during culture is not clear. However, it has been suggested that ES cell lines accumulate chromosomal abnormalities during culture (Longo et al., 1997; Rebuzzini et al., 2008). Rebuzzini et al., (2008) demonstrated that the frequency of euploid cells within an ES cell population decreases with an increasing number of passages. In this study 3 ES cell lines were karyotyped at 2 different passage numbers. At early passage numbers 35% of metaphase spreads from ES cell line 1 (passage 13), 49% from cell line 2 (passage 9), and 37% from cell line 3 (passage 7) were found to be aneuploid. These numbers increased to 61%, 68% and 79% at passages 31, 29 and 22, respectively (Rebuzzini et al., 2008). Those data are in accordance with another study which reported that ES cells up to passage 15 show around 40% aneuploidy which increases to 70-80% by passage 25 (Longo et al., 1997). Taken together those data suggest that on average, 40% of ES cells within a population are aneuploidy between passages 7-15. This number seems to be reasonable as another resent study observes similar numbers of aneuploidy in wild type ES cell lines (33%) although it is not clear at which passage number those cells were analysed (Fütterer et al., 2011). All ES cell lines analyzed by karyotyping during the course of this thesis had been cultured for 10-12 passages at the time of examination. Figure 4.7 shows that the aneuploidy levels in the control cell lines were around 19% which is about half the amount of aneuploidy observed by Longo et al., (1997) and Rebuzzini et al., (2008). The reason for the difference in aneuploidy levels between the control cell lines and the published data is not clear but it is possible that this discrepancy is, as mentioned above, generated by a certain culture environment and/or mouse strains used to generate the ES cell lines. Occurrence of aneuploidy in *Tex19.1*$^{-/-}$ ES cell lines is, at 74%, significantly higher than in the control cells (χ^2-test $p < 0.0005$). Importantly, all cell lines were analysed at similar passage numbers. This suggests that *Tex19.1* could be either involved in faithful chromosome segregation, or that gain or loss of certain chromosomes might compensate for the *Tex19.1* mutation, and therefore confers a growth advantage of aneuploid over euploid *Tex19.1*$^{-/-}$ cells which would result in aneuploid cells taking over the population. As the data presented here merely represent an endpoint, further work

would be required to understand the mechanisms that cause aneuploidy in *Tex19.1*$^{-/-}$ ES cells. It would be useful to establish euploid clones from each cell line and monitor their karyotype at several passages in order to confirm the initial result presented in this thesis. Ideally, this experiment would be carried out with conditional *Tex19.1*$^{-/-}$ ES cells which would avoid genotypic variation between cell lines as variability might be present in cell lines that have been generated from different embryos on a mixed genetic background. Furthermore, analysis of chromosomes aligned on the mitotic spindle could provide better insights into this phenotype. If an increase of lagging chromosomes could be observed in the *Tex19.1*$^{-/-}$ mutants compared to controls it would provide evidence that *Tex19.1* might be actively involved in faithful chromosome segregation in ES cells.

4.3.5 A role for *Tex19.1* during embryonic development

Results in this chapter have shown that *Tex19.1*$^{-/-}$ mice are born at a reduced rate from the expected Mendelian ratio. I have provided evidence that this reduction is exacerbated in embryos born from lactating mothers and that under those circumstances, females are stronger affected than males. Furthermore, I have extended our knowledge about the expression pattern of Tex19.1 and shown that, in addition to the previously reported expression in ES cells and the pre-implantation embryo, *Tex19.1* is expressed in the pluripotent cells of the early-post implantation embryo as well was in the extraembryonic cell lineages. Analysis of *Tex19.1*$^{-/-}$ and control ES cells failed to identify retrotransposon upregulation in ES cells in the absence of Tex19.1, but showed that *Tex19.1*$^{-/-}$ ES cells accumulate aneuploidy. If the aneuploidy is caused by chromosome segregation defects in the absence of Tex19.1 in ES cells, and if this could provide an explanation for the reduced viability of *Tex19.1*$^{-/-}$ homozygotes requires further investigation. The idea that *Tex19.1* might be involved in faithful chromosome segregation is further discussed in chapter 6.

Chapter 5: Characterisation of the role of *Tex19.1* during female germline development

5.1 Introduction

The main focus of my PhD project was to analyse the function of Tex19.1 in the female germline. As mentioned under 4.1, *Tex19.1* expression has been identified in migrating PGCs from 9.5dpc onwards by qRT-PCR (Hackett et al., under revision). RNA expression was also confirmed by qRT-PCR in gonads of both sexes at 10.5dpc and 13.5dpc or 13.5dpc through to adulthood by Hackett et al., (under revision) and Kuntz et al., (2008), respectively. In addition, Kuntz and colleagues demonstrated expression of *Tex19.1* in male and female gonads by in situ hybridization (ISH) which showed expression between 12.5-14.5 dpc in the males and 12.5-14.5 dpc in the females (Kuntz et al., 2008). Analysis of female gonads by ISH at 15.5dpc showed no expression of *Tex19.1*, however, this is in contrast with RT-PCR data presented in the same study which suggest *Tex19.1* RNA expression in female gonads at this stage (Kuntz et al., 2008). It is not clear which of the data is correct but it seems likely that *Tex19.1* RNA is expressed at 15.5dpc in the female germline as for the ISH a positive control is lacking, and *Tex19.1* RNA expression was also observed at 16.5 and 18.5dpc until after birth by RT-PCR (Kuntz et al., 2008). The stages at which growing oocytes express *Tex19.1* are less clear, but previous RT-PCR data and the EST expression profile from the unigene database (NCBI) suggest that *Tex19.1* is expressed in the adult ovary (Childs, 2006). Furthermore, *Tex19.1* RNA expression has been reported in GV, MI and MII oocytes by microarray and qRT-PCR (Chen at al., 2011).

Previous observations from Rupert Öllinger showed that in contrast to *Tex19.1*$^{-/-}$ knockout males, which are largely infertile, *Tex19.1*$^{-/-}$ knockout females are fertile on a mixed genetic background (129/Ola x CD1) but have ~ 50% reduction in litter size compared to *Tex19.1*$^{+/-}$ females (Öllinger et al., 2008). In this chapter I take those initial observations further and confirm fertility defects of *Tex19.1*$^{-/-}$ females on a C57/BL6 inbred background. I also investigate if the reduction in litter size observed in *Tex19.1*$^{-/-}$

females results from perturbations in the progression through meiosis and how this phenotype compares with the defects observed in *Tex19.1*$^{-/-}$ males.

5.2 Results

5.2.1 *Tex19.1*$^{-/-}$ females have reduced fertility

Here, I analyze the functional role(s) of *Tex19.1* in female germ cell development. In order to test if *Tex19.1*$^{-/-}$ females show fertility defects on an inbred C57/BL6 background, seven *Tex19.1*$^{-/-}$ and seven control females were mated with wild-type males. All females used in this study were under 6 months of age. In contrast to *Tex19.1*$^{-/-}$ males, *Tex19.1*$^{-/-}$ females were fertile and generated healthy offspring (figure 5.1). However, *Tex19.1*-deficient females exhibit a strong (~33%) reduction in litter size compared to control females (figure 5.1). On average *Tex19.1*$^{-/-}$ females generated 5.46 ± 3.95 offspring, compared with 8.18 ± 2.4 offspring for their heterozygote siblings (Mann-Whitney U test, $p<0.05$, n = 25 litters each).

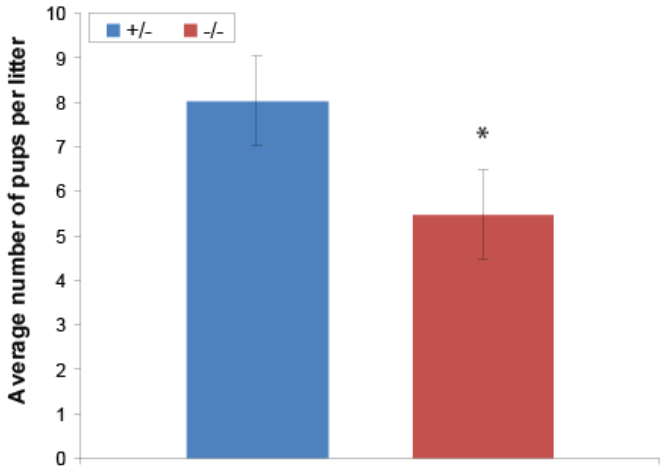


Figure 5.1 Average litter size observed in *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ females when mated with wild type males. Analysis of pups born from seven *Tex19.1*$^{-/-}$ and seven *Tex19.1*$^{+/-}$ females shows a ~ 33% reduction in the numbers of pups from *Tex19.1*$^{-/-}$ mothers compared to controls (Mann-Whitney U test, $p<0.05$). A total of 25 litters was analysed for each genotype. Graphs represent average number of pups ± standard error for all litters born per genotype.

5.2.2 *Tex19.1*$^{-/-}$ females transmit aneuploidies to their offspring

The reduction in litter size observed in *Tex19.1*$^{-/-}$ deficient females could be due to ovarian failures, resulting in functional oocyte loss. However, a comparison of ovarian morphology in Tex19.1$^{-/-}$ deficient and control females indicated that follicular development is not affected in the absence of *Tex19.1* (figure 5.2).

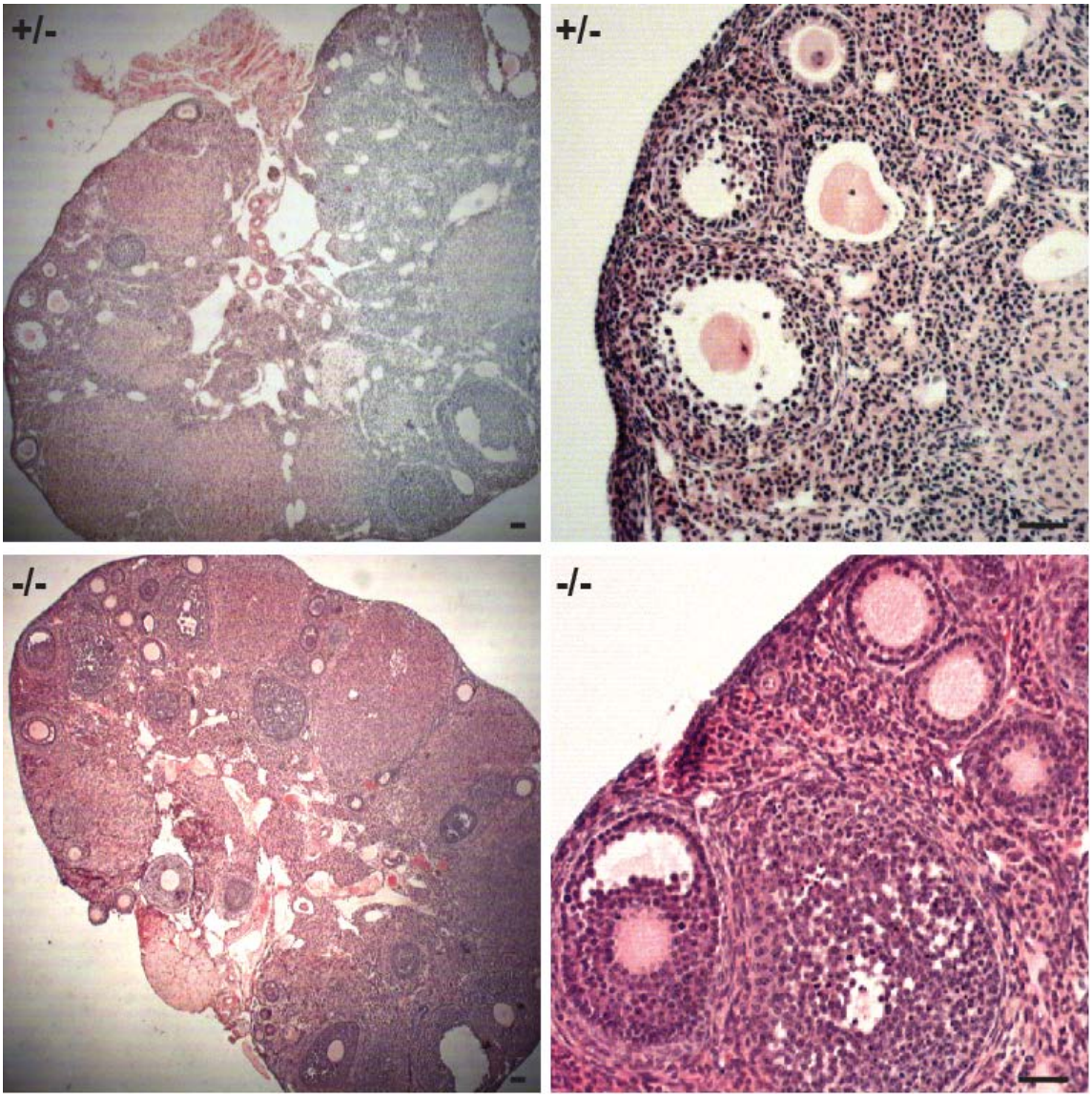


Figure 5.2 Ovary histology of *Tex19.1*$^{-/-}$ and control females. Haematoxylin and eosin stained ovary sections from *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ females. Primordial follicles and all stages of growing follicles are present in *Tex19.1*$^{-/-}$ ovaries. Scale bars are 100μm

This observation is supported by the fact that *Tex19.1*$^{-/-}$ females ovulate on average 9.5 ± 1.4 oocytes which is comparable to the 9.3 ± 1 oocytes ovulated by *Tex19.1*$^{+/-}$ females (two-tailed t-test, p>0.5) (figure 5.3). The data presented in figure 5.3 were obtained by counting the numbers of zygotes obtained from 10 *Tex19.1*$^{-/-}$ and 10 *Tex19.1*$^{+/-}$ females which had been mated with wild type males. Taken together, this suggests that the reduction in litter size, observed in *Tex19.1*$^{-/-}$ females, is likely caused by a reduction in oocyte quality rather than oocyte number.

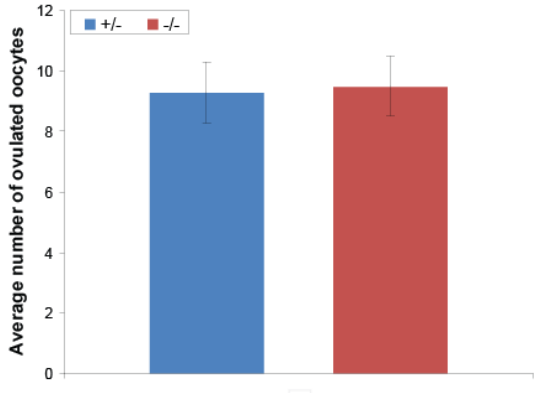


Figure 5.3 Average number of ovulated oocytes in *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ females. Analysis of zygotes from 10 *Tex19.1*$^{-/-}$ and 10 *Tex19.1*$^{+/-}$ females shows that *Tex19.1*$^{-/-}$ females ovulate the same number of oocytes as controls (two-tailed t-test, p>0.5). Graphs represent mean ± standard error for biological replicates.

Based on the data presented in figure 5.3, the germ cell restricted expression of *Tex19.1* in the adult mouse and the chromosome segregation defect observed in *Tex19.1*$^{-/-}$ males. we hypothesized that the reduction in litter size in *Tex19.1*$^{-/-}$ females could be due to generation of aneuploid oocytes resulting from chromosomal segregation errors. To test this we karyotyped one-cell zygotes (0.5dpc) from *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ females which revealed ~ 30% increase in chromosomal aberrations in zygotes derived from *Tex19.1*$^{-/-}$ compared to control females (two-tailed Fisher's exact-test, p<0.004) (figure 5.4). Of 56 zygotes obtained from *Tex19.1*$^{-/-}$ females, 64.5% were found to be euploid, 21.25% hypoploid and 14.25 % hyperploid. In contrast, out of 58 zygotes derived from *Tex19.1*$^{+/-}$ females 93% were euploid and 7% were hypoploid. The hypoploidy observed in oocytes derived from *Tex19.1*$^{+/-}$ females is likely, at least to some extent, caused by

technical difficulties during preparation of the chromosome spreads. Taken together those data suggest that the reduction in litter size observed in *Tex19.1*$^{-/-}$ females is caused by transmission of aneuploidy to their offspring rather than a reduction in the number of oocytes ovulated.

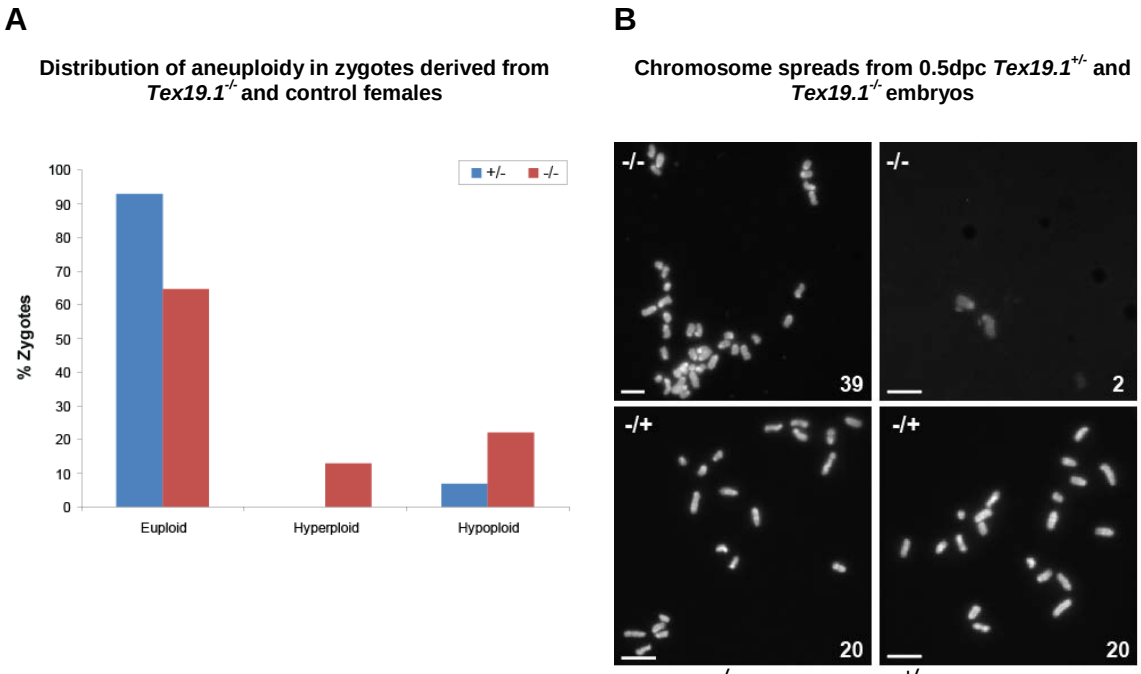


Figure 5.4 Average karyotype of zygotes from *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ females. (A) The diagrams show distribution of euploid, hyperploid and hypoploid cells derived from *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ females. Analysis of zygotes (0.5dpc) from 11 *Tex19.1*$^{-/-}$ and 13 *Tex19.1*$^{+/-}$ females, which had been mated to wild type males, identified a ~ 30% increase in the number of aneuploid zygotes compared to controls (for statistical analysis numbers of hyperploid and hypoploid cells were pooled; two-tailed Fisher's exact-test, $p<0.0004$). A total of 58 zygotes from 13 *Tex19.1*$^{+/-}$ and 56 zygotes from 11 *Tex19.1*$^{-/-}$ females were analysed for this experiment. (B) Spreads of metaphase chromosomes from 0.5dpc 1-cell embryos obtained from *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ embryos. Both pictures for *Tex19.1*$^{-/-}$ (41 chromosomes) and *Tex19.1*$^{+/-}$ zygote (40 chromosomes) belong to the same embryo. Scale bars are 30µm.

As mentioned above, the chromosomal abnormalities observed in one-cell zygotes from *Tex19.1*-deficient oocytes are possibly caused by chromosome segregation defects occurring during the progression through meiosis. It has been shown that the severe subfertility observed in *Tex19.1*$^{-/-}$ males is caused by germ cell apoptosis resulting from defects in chromosome synapsis during prophase I and defects in establishing or maintaining the bivalent at metaphase I (Öllinger et al., 2008). We postulated that

aneuploidy in offspring from *Tex19.1*$^{-/-}$ females might result from similar meiotic problems during oogenesis as observed in spermatogenesis in the absence of Tex19.1. To test this, I performed chromosome spreads on oocytes isolated from 17.5dpc fetal ovary and analysed chromosome synapsis. However, in contrast to *Tex19.1*$^{-/-}$ spermatocytes, *Tex19.1*$^{-/-}$deficient oocytes appear to complete chromosome synapsis normally (figure 5.5). Of Sycp3/Sycp1 stained spreads, 69% and 72% featured complete synapsis at pachytene, 4% and 5% were zygotene, 19% and 16% were diplotene and 2% and 3% showed asynapsis at pachytene in 244 *Tex19.1*$^{-/-}$ and 211 *Tex19.1*$^{+/-}$ oocytes, respectively. In addition to defects in chromosome synapsis during prophase I, *Tex19.1*$^{-/-}$ males show high rates of univalent chromosomes at metaphase I suggesting that formation or maintenance of bivalents at metaphase I might be impaired in the male germline in the absence of Tex19.1 (Öllinger et al., 2008). In order to test if a similar defect might result in the aneuploidy rates observed in embryos obtained from *Tex19.1*$^{-/-}$ females, chromosome spreads of pro-metaphase or metaphase I oocytes were performed. Of 59 and 71 pro-metaphase I or metaphase I spreads, 98.3% and 98.6% featured 20 bivalent chromosomes and 1.7% and 1.4% contained univalent chromosomes in *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ oocytes, respectively (figure 5.6). The data presented here suggest that *Tex19.1*$^{-/-}$ females complete chromosome synapsis and formation of the bivalent normally.

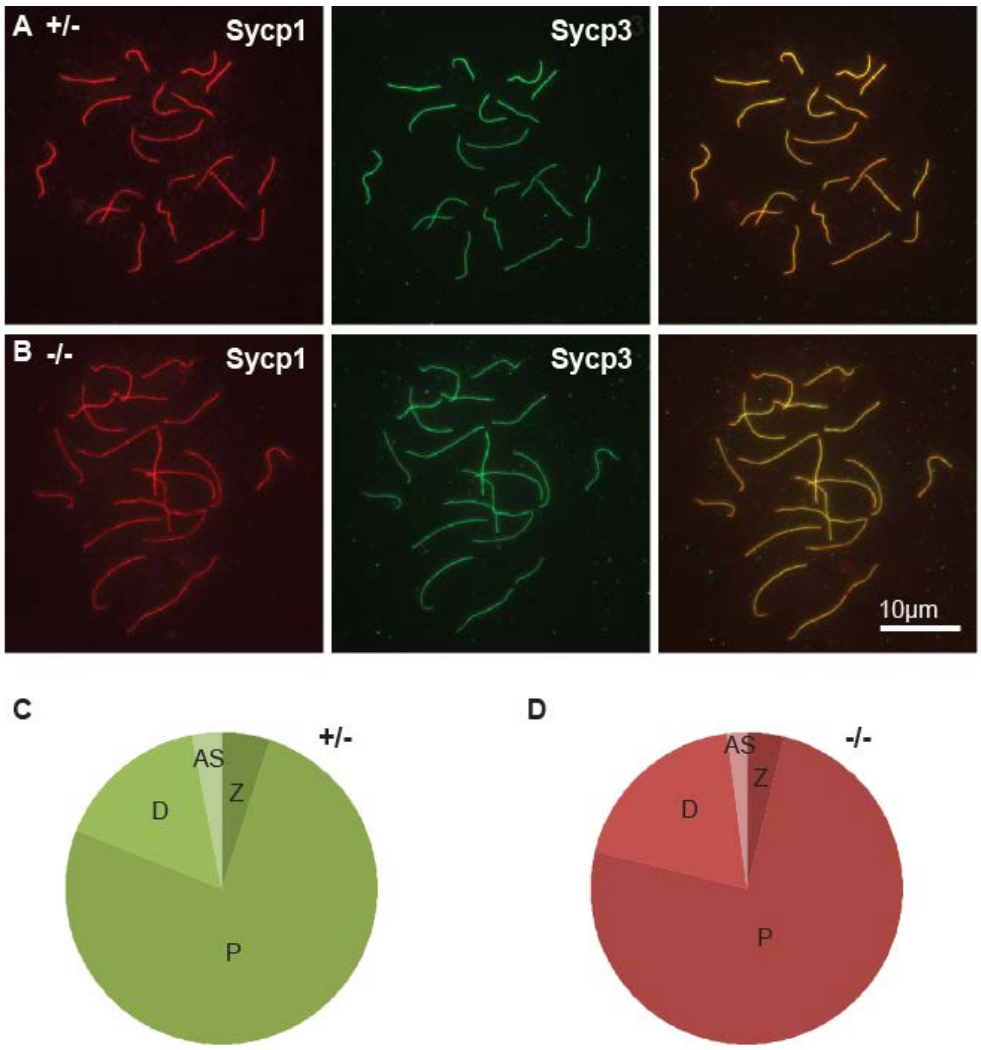
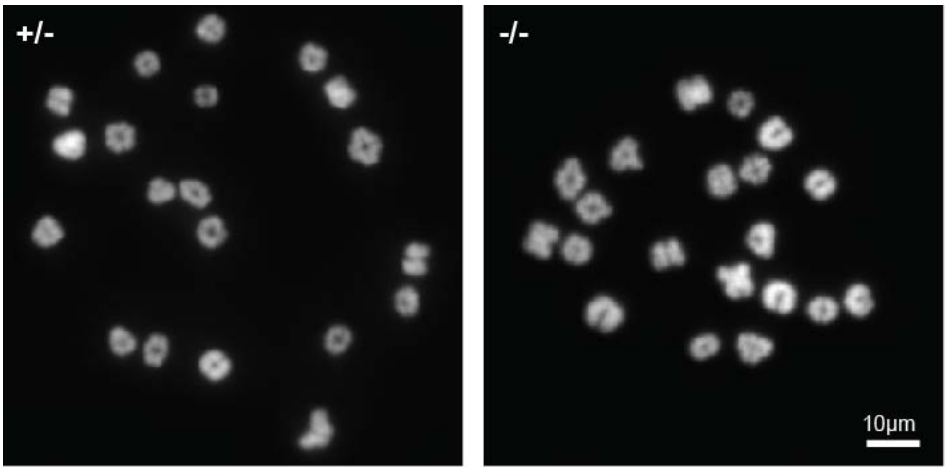


Figure 5.5: Chromosome synapsis in *Tex19.1*$^{-/-}$ null oocytes and control oocytes. (A and B) Synapsis of homologous chromosomes in meiotic cells; Sycp3 (green) is found on both synapsed and unsynapsed chromosomes during early meiotic prophase whereas Sycp1 (red) is only detected on synapsed chromosomes. *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ pachytene cells show complete synapsis of all chromosomes. (C and D) Distribution of meiotic stages in Sycp3/Sycp1-stained spreads (χ^2-test $p > 0.1$). A total of 244 chromosome spreads from five *Tex19.1*$^{-/-}$ animals and a total of 211 chromosome spreads from five *Tex19.1*$^{+/-}$ animals were scored (Z is zygotene, P is pachytene, D is diplotene, AS is asynapsis).

A

(Pro)-metaphase I chromosome spreads



B

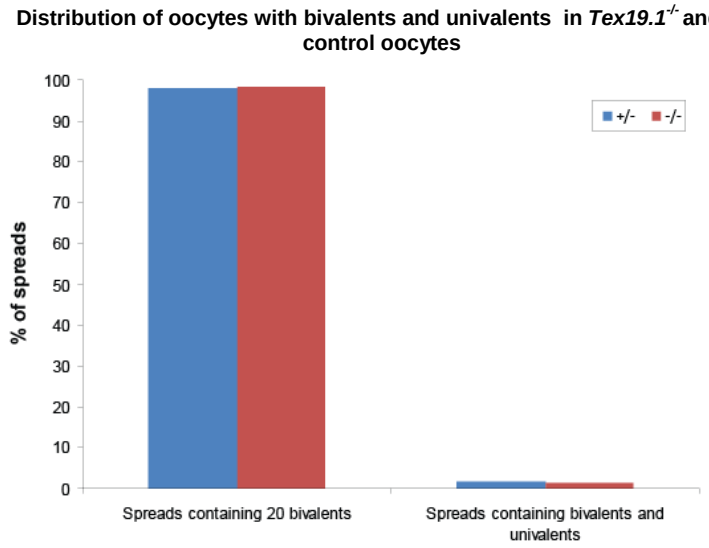


Figure 5.6: Meiotic chromosome behavior in metaphase I spreads derived from *Tex19.1*$^{-/-}$ and control oocytes. (A) Bivalent formation of chromosomes in meiotic cells. (B) Distribution of spreads with 20 bivalents and spreads containing univalents in 59 *Tex19.1*$^{+/-}$ oocytes and 71 *Tex19.1*$^{-/-}$ oocytes (two-tailed Fisher's exact-test, p>0.5)

This result is again in contrast with the meiotic defects observed during spermatogenesis in the absence of *Tex19.1* (Öllinger et al., 2008). However, since *Tex19.1*$^{-/-}$ females transmit aneuploidies to their offspring we postulated that defects in chromosome segregation must be evident in metaphase II or anaphase II oocytes. To test this I

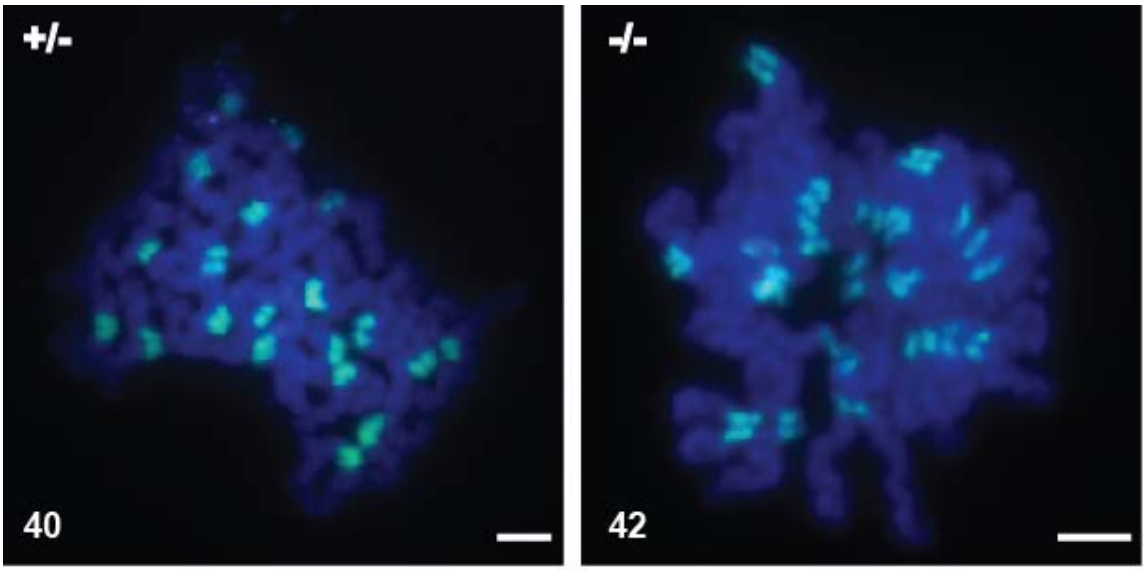
performed chromosome spreads from *Tex19.1*$^{-/-}$ and control oocytes. In order to do so, females were treated with PMS and HCG to induce superovulation. As mentioned in chapter 1, ovulated mammalian oocytes remain in metaphase II until entry of the sperm induces resumption of meiosis. Therefore spreads to be examined for chromosome behavior during metaphase II were collected and analysed immediately after superovulation. The chromosomes were stained with DNA FISH for major satellites, very large arrays of tandemly repeating, non-coding DNA which constitute the main component of functional centromeres, in order to facilitate the chromosome counts. Figure 5.7 shows that aneuploidies increase significantly by ~ 30% in metaphase II oocytes from *Tex19.1*$^{-/-}$ females compared to controls (two-tailed Fisher's exact-test, $p<0.004$). For each genotype, 48 oocytes were analysed. This suggests that in *Tex19.1*$^{-/-}$ females chromosome mis-segregation occurs during the metaphase I to anaphase I transition.

To further strengthen our hypothesis that *Tex19.1*$^{-/-}$ females transmit aneuploidy to their offspring, I analysed chromosome spreads prepared from anaphase II oocytes. Activation events, that trigger resumption and completion of meiosis in metaphase II arrested oocytes, are induced by repetitive intracellular Ca^{2+} oscillations caused by oocyte fertilization by the sperm. During this project I cultured metaphase II oocytes in Ca^{2+} chelated medium and induced meiotic activation by supplementing the media with Sr^{2+} as described (Kishigami and Wakayama, 2007). Firstly, I established conditions and timings that would allow me to analyse oocytes at anaphase II. Figure 5.8 shows that under these conditions (see chapter 2) after 30 minutes 87.5% of oocytes were still found to be in metaphase II, after one hour 55% of oocytes were found to be still at metaphase II and after two hours of culture 80% of oocytes were found to be in anaphase II. After 4 hours the majority of oocytes (83%) had progressed beyond meiosis. Accordingly, *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ oocytes were analysed for chromosome behaviour during anaphase II after two hours of culture. In accordance with the breeding, zygote and MII data, *Tex19.1*$^{-/-}$ anaphase II oocytes showed ~ 34% increase in the occurrence of aneuploidy compared to controls (two-tailed Fisher's exact-test $p<0.005$) (figure 5.9). A

total of 35 oocytes from *Tex19.1*$^{-/-}$ and 24 oocytes from *Tex19.1*$^{+/-}$ females were analysed. Taken together analysis of 0.5dpc embryos, metaphase II and anaphase II oocytes suggest that around 30% of *Tex19.1*$^{-/-}$ oocytes are aneuploid. The abnormal karyotype seems to arise from a defect in chromosome segregation during the metaphase I to anaphase I transition which is then transmitted to subsequent generations resulting in embryonic lethality.

Further analysis of the chromosome aberrations found in metaphase II oocytes, anaphase II oocytes and zygotes derived from *Tex19.1*$^{-/-}$ females showed that the observed aneuploidies resulted from gain or loss both of even and odd numbers of chromosomes (figure 5.10A and B). This suggests that chromosome mis-segregation in the absence of *Tex19.1* is caused by additional or other defects than those that cause classical non-disjunction as classical non-disjunction leads to loss/gain of 2, 4, 6 etc chromosomes. In contrast premature separation of sister chromatids could lead to loss/gain of 1, 2, 3, 4, 5 etc chromosomes generating oocytes with even and odd aneuploidies (see also figure 5.10C).

A



B

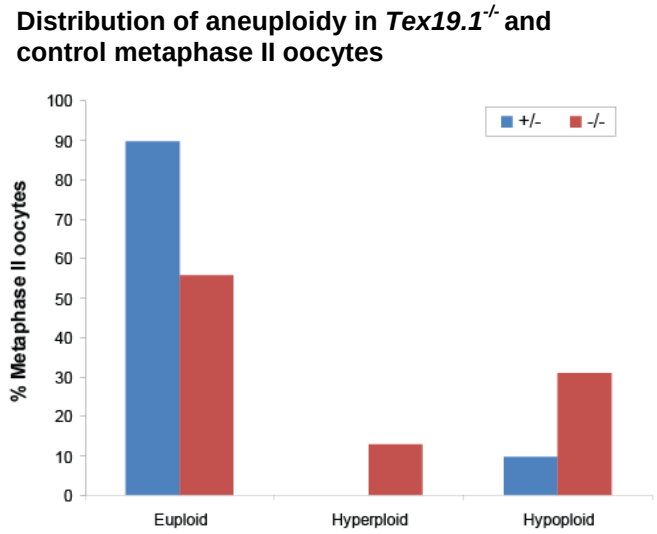


Figure 5.7: Meiotic chromosome behavior in metaphase II spreads derived from *Tex19.1*$^{-/-}$ oocytes and control oocytes. (A) Metaphase II chromosomes stained with DAPI for DNA (Blue) and DNA FISH for major satellites (green). (B) Distribution of euploid, hyperploid and hypoploid oocytes among spreads prepared from 48 *Tex19.1*$^{+/-}$ oocytes and 48 *Tex19.1*$^{-/-}$ oocytes. Aneuploidy levels in *Tex19.1*$^{-/-}$ metaphase II oocytes were found to be increased by ~ 30% compared to controls (for statistical analysis numbers of hyperploid and hypoploid cells were pooled; two-tailed Fisher's exact-test, $p<0.004$). Scale bars are 10µm.

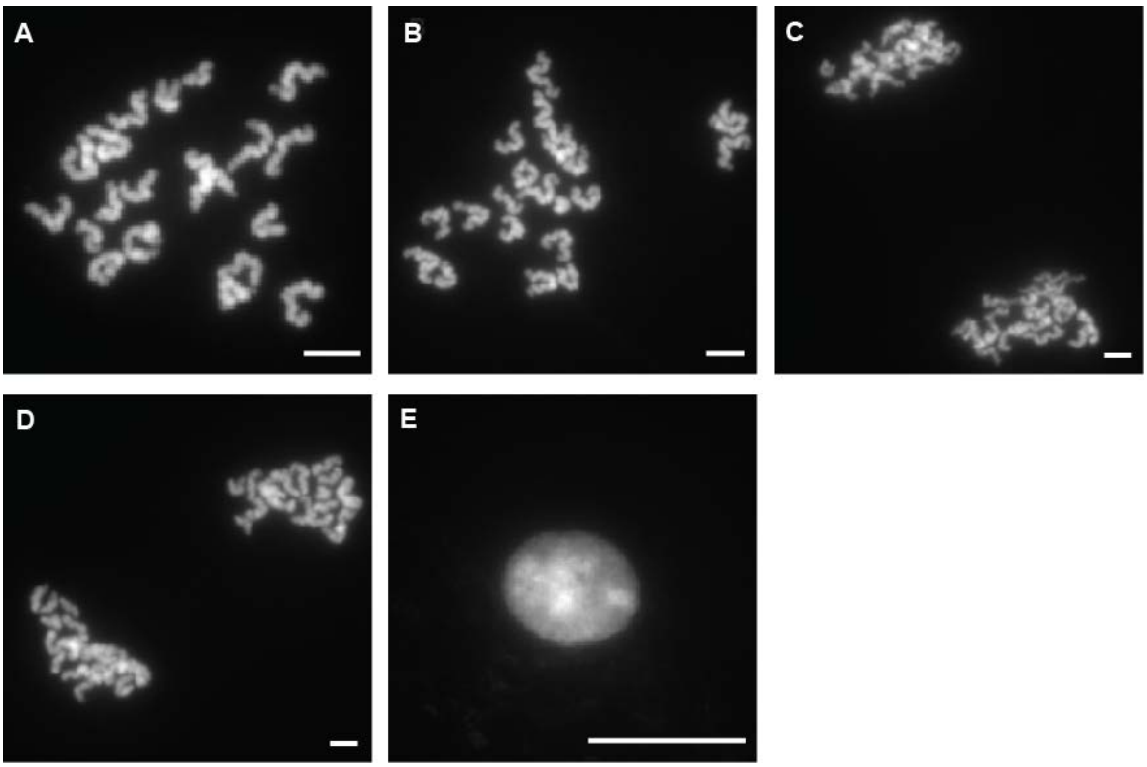
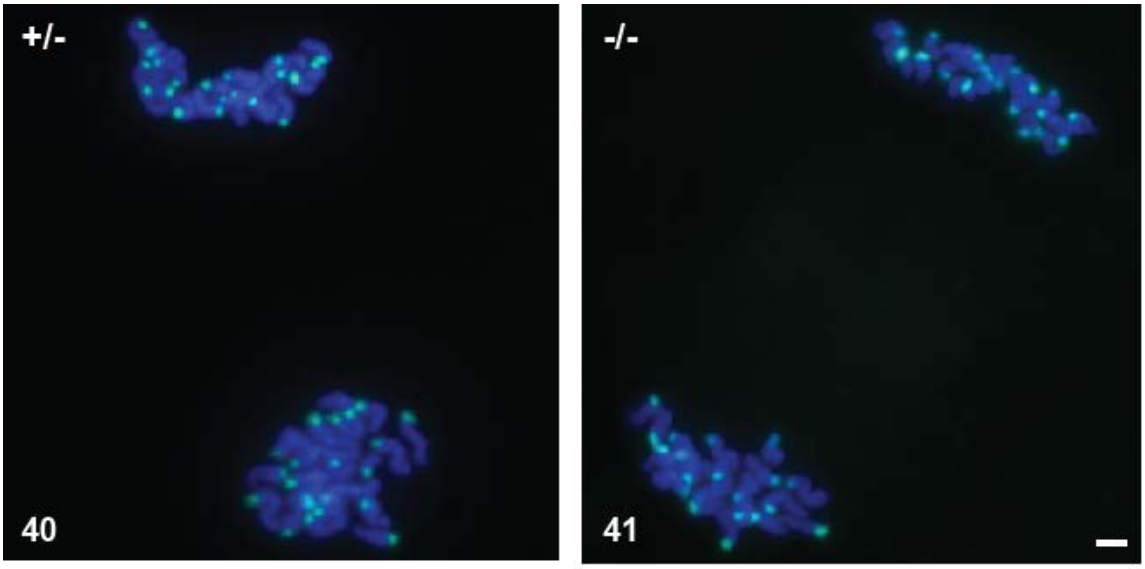


Figure 5.8: Metaphase II – Anaphase II transition in wild type mouse oocytes can be induced in culture. (A) Ten out of 10 oocytes collected and analysed immediately after superovulation (0h of culture) were found to be in metaphase II. (B) After 1/2h of culture 7 out of 8 oocytes were found to be in metaphase II, 1 was found to be in anaphase II. (C) After 1h of culture 6 out of 11 oocytes were found to be in anaphase II, 5 out of 11 oocytes had remained in metaphase II. (D) After 2h of culture 8 out of 10 oocytes were found in anaphase II, 2 out of 10 oocytes were found to be in metaphase II. (E) Analysis of oocytes after 4h of culture showed that 5 out of 6 oocytes were found in a state which we classified as post-anaphase II and 1 out of 6 had remained in metaphase II. Scale bars are 10μm.

A



B

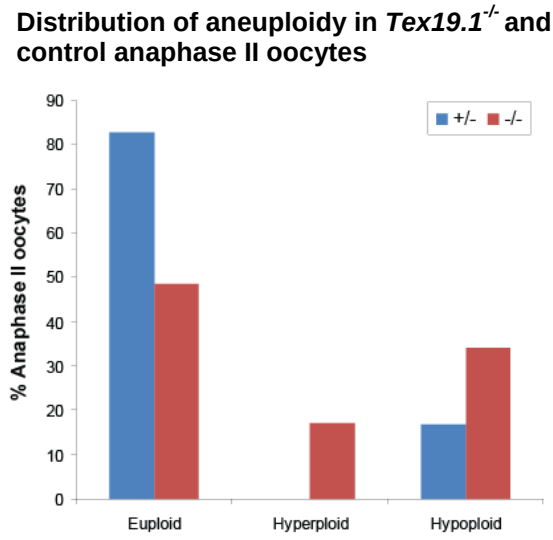
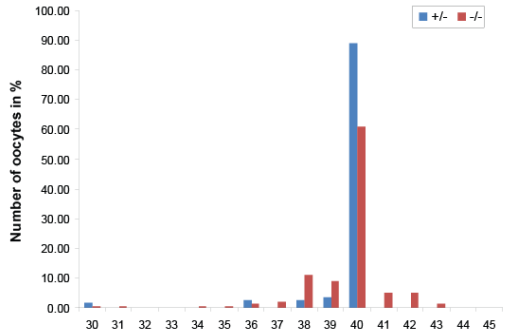


Figure 5.9: Meiotic chromosome behavior in anaphase II spreads derived from *Tex19.1*$^{-/-}$ oocytes and control oocytes. (A) Anaphase II chromosomes stained with DAPI for DNA (Blue) and DNA FISH for major satellites (green). (B) Distribution of euploid, hyperploid and hypoploid oocytes among spreads prepared from 24 *Tex19.1*$^{+/-}$ oocytes and 35 *Tex19.1*$^{-/-}$ oocytes. Aneuploidy levels in *Tex19.1*$^{-/-}$ anaphase II oocytes were found to be increased by ~ 34% compared to controls (for statistical analysis numbers of hyperploid and hypoploid cells were pooled; two-tailed Fisher's exact-test, $p<0.004$). Scale bars are 10µm.

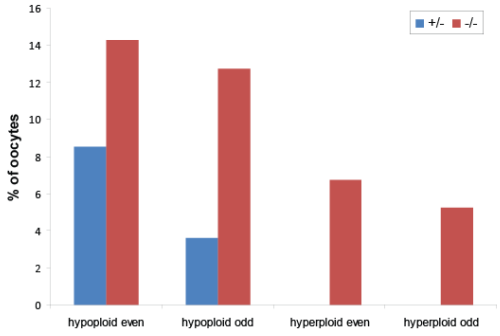
A

Karyotypes observed in oocytes and embryos derived from *Tex19.1*$^{-/-}$ and control females



B

Distribution of euploidy, hyperploidy and hypoploidy in oocytes and embryos derived from *Tex19.1*$^{-/-}$ and control females



C

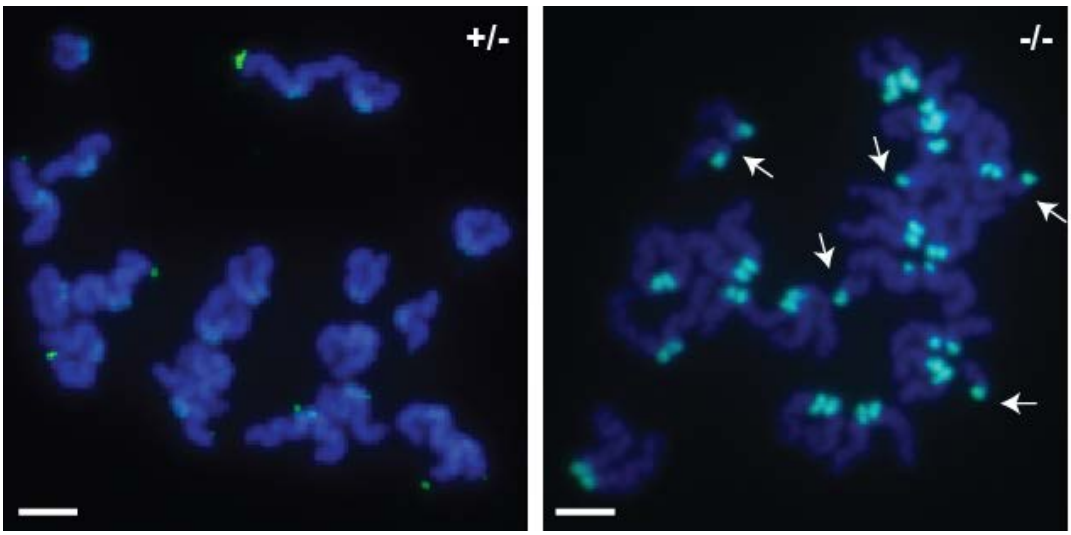


Figure 5.10: Karyotypes and distribution of aneuploidy in *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ derived from metaphase II and anaphase II oocytes and zygotes. (A) Distribution of chromosome numbers in metaphase II and anaphase II oocytes and one-cell zygotes. (B) Percentage of oocytes and zygotes with hyperploidy and hypoploidy arising from even and odd numbers of chromosomes. (A and B) The results from 130 oocytes and zygotes derived from *Tex19.1*$^{+/-}$ and 139 derived from *Tex19.1*$^{-/-}$ females are pooled and summarized in this analysis. (C) Metaphase II chromosomes from *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ oocytes stained with DAPI for DNA (Blue) and DNA FISH for major satellites (green). Arrows indicate single chromatids. Scale bars are 10μm.

5.3 Discussion

In this chapter I aimed to gain a better understanding of the role of *Tex19.1* during female meiosis. I show that in the absence of *Tex19.1*, females have reduced litter size and that this is likely caused by transmission of an abnormal karyotype to the embryo by the oocyte. I demonstrate that in *Tex19.1*$^{-/-}$ oocytes, in contrast to *Tex19.1*$^{-/-}$ spermatocytes, synapsis and formation and/or maintenance of the bivalent of homologous chromosomes appears to occur normally but that high levels of aneuploidy become evident during the second round of meiosis. *Tex19.1* is thus linked to inherited aneuploidy in female germ cells and *Tex19.1*$^{-/-}$ mice could provide a valuable model for studying mechanisms that lead to chromosome mis-segregation during oogenesis.

5.3.1 *Tex19.1*$^{-/-}$ females have reduced fertility

Breeding analysis of *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ females with wild type males revealed a ~ 33% reduction in litter size in *Tex19.1*$^{-/-}$ females compared to controls (figure 5.1). Causes for female fertility defects can be various: inappropriate hormone levels or responses to hormones, sexually transmitted diseases, being underweight or overweight, inability to implant the embryo/s, reduction in oocyte number or oocyte quality can all have a negative effect on female fertility, to name just a few (Steinberger et al., 1979; Hassold and Hunt, 2001; Filis et al., 2009; Akande et al., 2010; Jagarlamudi et al., 2010; Souter et al., 2011). As mentioned before, *Tex19.1* expression seems to be restricted to the germ cells in the adult mouse and therefore we reasoned that the reduction in litter size observed in *Tex19.1*$^{-/-}$ females might be caused by a reduction in oocyte number or oocyte quality. Premature reduction in the number of oocytes affects about 1% of human females and is caused by premature ovarian failure (POF), a complex disorder characterized by a deficiency of oocytes resulting in diminished fertility and shortened reproductive lifespan (Jagarlamudi et al., 2010). As described in chapter 1, defects in ovarian differentiation, primordial follicle formation, follicular growth, atresia or

alterations in ovulation can result in a shortage of oocytes. Many mutant mouse models, with mutations in genes that affect PGC development and germ cell proliferation or survival as well as in genes that cause severe defects in meiosis when mutated, exhibit a lack of functional oocytes (Luoh et al., 1997; Ruggiu et al., 1997; Di Giacomo et al., 2005). However, in *Tex19.1*$^{-/-}$ females the reduction in litter size appears to be caused by a reduction in oocyte quality rather than oocyte number as *Tex19.1*$^{-/-}$ females ovulate the same number of oocytes as controls. In addition, analysis of one-cell zygotes from *Tex19.1*$^{-/-}$ and control females revealed a ~ 30% increase in aneuploidy in embryos derived from *Tex19.1*-deficient mothers (figures 5.2 and 5.3). All females used in this study were mated between 6 weeks and 6 month of age. Reduction in litter size in *Tex19.1*$^{-/-}$ females was evident from 6 weeks onwards and did not increase in ageing females up to 6 month of age.

5.3.2 *Tex19.1*$^{-/-}$ oocytes show chromosomal abnormalities in meiosis II

Based on the aneuploidy identified in zygotes derived from *Tex19.1*$^{-/-}$ females, we proposed that the reduction in litter size observed in *Tex19.1*$^{-/-}$ females is likely caused by chromosome segregation defects during meiosis, resulting in aneuploidy and therewith leading to embryo death. As mentioned in the previous sections *Tex19.1*$^{-/-}$ males show defects in homologue chromosome synapsis and formation or maintenance of bivalent chromosomes at metaphase I. These defects trigger apoptosis in spermatocytes which prevents chromosomal abnormalities being transmitted to subsequent generations but also leads to male infertility (Öllinger et al., 2008). In female germ cells some mutations in meiotic genes that cause spermatocyte death are tolerated resulting in the formation of aneuploid oocytes and reduced fertility due to high rates of embryo lethality. For example *Sycp3*$^{-/-}$ mice, which show defects in synaptonemal complex formation, are characterised by male infertility but female subfertility (Yuan et al., 2000, Yuan et al., 2002). Subfertility in *Sycp3*$^{-/-}$ females appears to be caused by

transmission of high rates of aneuploidy to their offspring rather than by a reduction of oocyte number (Yuan et al., 2002). As described under 1.6.2 an abnormal karyotype results in growth and developmental retardation of the affected embryo and often lethality during gastrulation (Hernandez and Fisher, 1999; Lightfoot et al., 2005).

In *Tex19.1*$^{-/-}$ mice not only the fertility phenotype appears to display sexual dimorphism but also the meiotic stage at which chromosome mis-segregation seems to occur. In contrast to *Tex19.1*$^{-/-}$ males my data suggest that *Tex19.1*$^{-/-}$ females complete chromosome synapsis normally (figure 5.4) (Öllinger et al., 2008). Similarly, bivalent formation appears to be unaffected in *Tex19.1*$^{-/-}$ females (figure 5.5). However, I was able to show that aneuploidy is evident in around 30% of metaphase II *Tex19.1*$^{-/-}$ oocytes (figure 5.6). Chromosome spreads from anaphase II oocytes confirm the increase in aneuploidy in *Tex19.1*$^{-/-}$ females observed at metaphase II and in one-cell zygotes (figure 5.8). Aneuploidy levels at those three different stages were found independently of each other to be increased by around 30% in *Tex19.1*$^{-/-}$ females compared to controls (figures 5.2, 5.5 and 5.6). The incidence of aneuploidy in metaphase II oocytes, anaphase II oocytes and zygotes from *Tex19.1*$^{-/-}$ females is consistent with the ~ 30% reduction in litter size observed in *Tex19.1*$^{-/-}$ females. This supports the hypothesis that female subfertility in *Tex19.1*$^{-/-}$ mutants is caused by chromosome mis-segregation during meiosis. The fact that chromosomal abnormalities are evident in metaphase II oocytes suggests that the aneuploidy has been generated at or between metaphase I and/or anaphase I. A good experiment to support the theory that chromosome mis-segregation during the first meiotic division causes aneuploidy in *Tex19.1*$^{-/-}$ females would be real time imaging of oocytes injected with mCherry-tagged H2B mRNA to visualise the chromosomes. Other studies have successfully used this technique to show lagging chromosomes, which are an indication of chromosome segregation defects, in mouse oocytes during meiosis (Chiang et al., 2010; Jin et al., 2010; Lister et al., 2010). Bivalent formation in *Tex19.1*$^{-/-}$ oocytes was analysed after 3h of culture following germinal vesicle breakdown. It is likely that those chromosomes represent the pro-metaphase I or early metaphase I stage as it has been previously reported that in cultured mouse oocytes

metaphase I is first appearing after $4^{1/2}$ h. Anaphase I can be first observed after 9h of culture (Donahue, 1968). It would also be interesting to analyse *Tex19.1*$^{-/-}$ oocytes at the late metaphase I stage in order to examine if the aneuploidy observed later in meiosis could result from defects in maintaining the bivalent in the absence of *Tex19.1*. In the same context it would be useful to analyse cross-over formation in *Tex19.1*$^{-/-}$ females by staining of pachytene oocytes with Mlh1 antibody, as changes in levels or distribution of chiasmata have previously been reported to cause aneuploidy in oocytes (Yuan et al., 2002; Hodges et al., 2005). However, defects in cross-over formation or maintenance would cause classical non-disjunction as described in figure 1.7 resulting in oocytes with gain or loss of even numbers of chromosomes. In contrast, analysis of the collected metaphase II, anaphase II and zygote data showed that *Tex19.1*$^{-/-}$ oocytes are characterised by both even and odd hypo- and hyperploidies suggesting that chromosome mis-segregation in the absence of *Tex19.1* is caused by additional or other defects than those that cause classical non-disjunction (figure 5.9). Premature separation of sister chromatids could lead to loss/gain of 1, 2, 3, 4, 5 etc chromosomes generating oocytes with even and odd aneuploidies making it a candidate for the chromosome abnormalities observed in *Tex19.1*$^{-/-}$ females (figures 1.8 and 5.9).

5.3.3 How does *Tex19.1* promote faithful chromosome segregation in the female germline?

During the course of this thesis I was able to advance our understanding of the *Tex19.1*$^{-/-}$ female fertility phenotype. I showed that subfertility in *Tex19.1*$^{-/-}$ females is caused by transmission of aneuploidy to their offspring. The mechanism by which Tex19.1 promotes faithful chromosome segregation will require further investigation. It is currently unclear if the meiotic problems observed in *Tex19.1*$^{-/-}$ females correlate with upregulation of *MMERVK10C* or other retrotransposons. *MMERVK10C* RNA levels in 15.5dpc fetal ovary were not found to be up-regulated in *Tex19.1*$^{-/-}$ mutants by qRT-PCR

or in adult ovary by ISH (Judith Reichmann, MRes project, 2008). However, *MMERVK10C* appears to be upregulated specifically at the pachytene stage in *Tex19.1*$^{-/-}$ spermatocytes and accordingly it would be useful to repeat this experiment on 18.5dpc fetal ovary as the majority of oocytes should be in pachytene at this stage. If the upregulation of retrotransposons in *Tex19.1*$^{-/-}$ testes occurs at the protein level, as speculated in chapter 3, upregulation of *MMERVK10C* at the RNA level might just be a down-stream consequence which might not occur in *Tex19.1*$^{-/-}$ females. Different timings of retrotransposon regulation and/or other male/female germ cell specific mechanisms for retrotransposon suppression could result in the sexually dimorphic phenotype observed in *Tex19.1*$^{-/-}$ mutants. Alternatively, *MMERVK10C* might be a target for Tex19.1 in male but not female germ cells. Other targets may or may not be shared between the sexes. It requires further investigation if the chromosome abnormalities during meiosis II, observed in *Tex19.1*$^{-/-}$ females, can also be found in the males, as due to the earlier meiotic defects present in *Tex19.1*$^{-/-}$ males most spermatocytes undergo apoptosis at the pachytene or metaphase I stage. As mentioned under 1.6.3 the characterisation of mouse knockout models has linked mutations in several meiotic pathways with chromosome mis-segregation. For example mutations in the synaptonemal complex protein Sycp3 or Mlh1, which is required for recombination foci to mature into chiasmata, result in defective recombination (Yuan et al., 2002; Kneitz et al., 2000). This will cause premature separation of homologous chromosomes by pro-metaphase I which then randomly segregate resulting in the formation of hypo- and hyperploidies with even numbers of chromosomes. However, aneuploid oocytes and zygotes from *Tex19.1*$^{-/-}$ females feature even and odd numbers of chromosomes indicating that the chromosome mis-segregation occurring between metaphase I and anaphase I in the absence of *Tex19.1* is possibly caused, at least in part, by premature separation of sister chromatids. As described in chapter 1 connection between sister chromatids during mitosis and meiosis is facilitated by cohesin protein complexes distributed along the chromosome arms as well as along the centromeres. During mitosis most of the cohesin complexes dissociate from the chromosome arms during prophase. At the onset of anaphase the remaining Scc1/Rad21 at chromosome arms and

centromeres is cleaved by seperase to allow sister chromatid segregation (Losada et al. 1998; Waizenegger et al. 2000). In contrast, meiotic chromosomes maintain sister chromatid cohesion until the onset of anaphase I, which contributes to the tethering of bivalents, in order to facilitate separation of homologue chromosomes (Revenkova et al. 2001; Eijpe et al. 2003. At the onset of anaphase I, degradation of Rec8, through seperase activity, releases cohesin from the chromosome arms (Kudo et al., 2006). Centromeric Rec8, protected by Sgol2, remains intact until anaphase II to allow attachment and movement of sisters to the same pole during the first meiotic division (Kudo et al., 2006). At the onset of anaphase II when sisters need to segregate Rec8 is cleaved to allow sister chromatid separation. *Sgol2*$^{-/-}$ mice are infertile and show a premature release of Rec8 cohesin complexes from anaphase I centromeres leading to complete loss of centromeric cohesion at this stage. This leads to single chromatids which can not be segregated appropriately at anaphase II resulting in the formation of aneuploid gametes that give rise to infertility (Lee et al., 2008; Llano et al., 2008). Defects in sister chromatid cohesin has also been observed in *Rad51c*-deficient, *Bub1*$^{+/-}$ and *Smc1β*$^{-/-}$ females (Kuznetsov et al., 2007; Leland et al., 2009). Kuznetsov and colleagues (2007) suggest that the phenotype observed in *Rad51c*-deficient is caused by a possible role of Rad51c in resolving the double Holliday junction (HJ). They propose that at anaphase I, when the microtubules have attached to the kinetochores of homologous chromosomes to pull them apart, unresolved double HJs may cause increased tension at the centromeres resulting in a disruption of sister chromatid cohesin (Kuznetsov et al., 2007. Data from XO mice, which show disrupted sister chromatid cohesion at the centromeres of the unpaired X chromosome at anaphase I, suggest that that the spindle force is indeed sufficient to cause premature sister chromatid separation. Consistent with this *Rad51c*-deficient metaphase II oocytes display also broken chromosomes (Kuznetsov et al., 2007). However, broken chromosomes were not observed in *Tex19.1*$^{-/-}$ oocytes suggesting that premature separation of sister chromatids might be caused by a premature loss rather than a disruption of sister chromatid cohesin by spindle forces. Female mice heterozygous for *Bub1*, a mitotic checkpoint kinase gene, transmit aneuploidy, resulting from premature sister chromatid separation during

the first meiotic division, to their offspring (Leland et al., 2009). It has been proposed that Bub1 plays a role in protecting centromere cohesion through recruitment of Sgol2 and other proteins and that this function is impaired in *Bub1*$^{+/-}$ oocytes (Leland et al., 2009). Premature sister chromatid separation during the first meiotic division is also observed in oocytes from older females deficient for the meiosis specific cohesin protein Smc1β (Revenkova et al., 2004). Furthermore, *Smc1β*$^{-/-}$ mutant oocytes show a destabilisation of chiasmata resulting in large proportion of homologous chromosomes being present as univalents at metaphase I (Hodges et al., 2005). Tex19.1 might fulfill a similar role to Bub1 or Smc1β by either protecting or facilitating the formation of sister chromatid cohesin. Loss of sister chromatid cohesin in *Tex19.1*$^{-/-}$ mutants might explain the phenotype described in this chapter,

Cohesin has been recently suggested to link mammalian female ageing and the increased occurrence of aneuploidy. The occurrence of chromosomal abnormalities was compared between oocytes of 2 and 14 month old mice which showed that cohesion and cohesion protector Sgol2 levels decline in aging oocytes (Lister et al., 2010). It was further shown that distance between sister chromatid centromeres was increased in aging oocytes suggesting an age related loss of centromere cohesion. Those observations correlated with chromosome segregation defects at anaphase I and mis-alignment of chromosomes at metaphase II (Lister et al., 2010). A study carried out by Chiang et al., (2010) came to similar conclusions. They also found that distance between centromeres was increased in females older than 16 month of age compared to 6–14 weeks old females correlating with decreased levels of Rec8 protein in oocytes derived from older females. This might predispose those oocytes to meiotic errors involving the premature separation of homologous and sister chromatids (Chiang et al., 2010). Furthermore, Hodges and colleagues (2005) proposed that a failure to maintain sister chromatid cohesin distal to the site of chiasmata formation might result in premature separation of homologous chromosomes. If sister chromatid cohesin indeed degenerates with increasing maternal age chromosomes with distal crossover events might be particularly vulnerable to an age related weakening of the cohesin mediated connection. Loss of distal cohesin will result

in the loss of the distal chiasmata and result in an achiasmate homologous pair that will segregate randomly at meiosis I (Jones, 2008). Figure 5.10 shows how weakened cohesion between sister chromatids can result in chromosome mis-segregation and demonstrates that the type of aneuploidies observed in *Tex19.1*$^{-/-}$ females would be consistent with a defect in sister chromatid cohesion. It would be valuable to perform chromosome spreads of late metaphase I *Tex19.1*$^{-/-}$ and control oocytes and analyse the distance between centromeres in order to gain further information whether or not sister chromatid cohesion might be affected in the absence of Tex19.1. Simultaneous antibody staining for Rec8 and Sgol2 should also shed further light on the matter. Possible reasons for sister chromatid cohesion impairment in *Tex19.1*$^{-/-}$ females and how these defects relate to human oocyte aging are discussed in chapter 6.

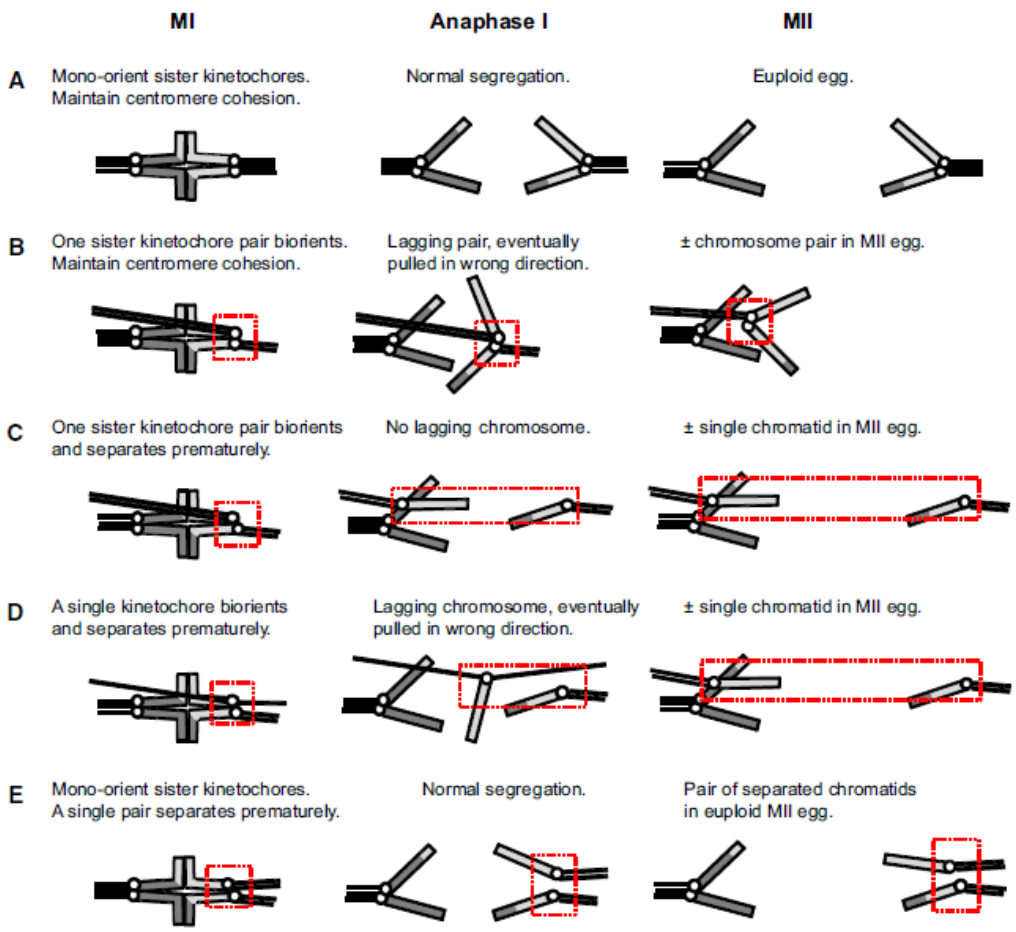


Figure 5.10: Proposed outcomes of weakened sister chromatid cohesion during meiosis. **(**A–E) Diagrams show proposed outcomes of normal (A) or weakened (B-E) sister chromatid cohesion in metaphase I (MI). Red boxes follow the behavior of a sister chromatid (or sister kinetochore) pair. The status of cohesion and sister kinetochore orientation on the metaphase I spindle determines how chromosomes segregate in anaphase I and subsequently the karyotype of the metaphase II oocyte. For prediction (E), sister chromatid separation is shown at metaphase I but could alternatively arise at a later time. Taken from Chiang et al., (2010).

Chapter 6: Discussion

The main aim of the research presented in this thesis, was to gain a better understanding of the *Tex19.1*$^{-/-}$ phenotype. Specific aims included examination of the basis of retrotransposon upregulation during spermatogenesis, characterisation of the function of *Tex19.1* during embryonic development, investigation of *Tex19.1* function during female germline development and consequently, an expansion of our knowledge concerning the role of *Tex19.1* during the germline cycle.

6.1 Retrotransposon upregulation, Tex19.1 and spermatogenesis

It was previously shown that *Tex19.1*$^{-/-}$ males are subfertile due to defects in chromosome segregation during meiosis, and that this phenotype correlates with RNA upregulation of the retrotransposon *MMERVK10C* (Öllinger et al., 2008). The phenomenon of retrotransposon mis-expression during meiosis has frequently been implicated in aberrant chromosome segregation (Bourc'his and Bestor, 2004; Kuramochi-Miyagawa et al., 2004; Carmell et al., 2007). To date however, it is not clear if retrotransposon upregulation is the cause or consequence of defects in meiotic progression but there is a growing number of mouse mutants which associate the two events (discussed in chapter 1). It is possible that retrotransposon upregulation is an indirect consequence of chromatin or gene expression changes caused by meiotic defects. However, *Mei1*$^{-/-}$ mice, which are characterised by disruptions in chromosome synapsis, show no upregulation of *LINE1* or *IAP* retrotransposon RNA (Libby et al., 2002; Carmell et al., 2007). This suggests that synapsis defects alone are not sufficient to cause mis-expression of retrotransposons. Despite the strong link between retrotransposon expression and defects in meiotic progression it is currently not clear how retrotransposons may cause problems in meiosis. Several possibilities are discussed below. Similar to *Dnmt3L*$^{-/-}$, *Miwi2*$^{-/-}$ and *Mili*$^{-/-}$ mutant males, which also show retrotransposon upregulation, *Tex19.1*$^{-/-}$ knockout spermatocytes are able to initiate meiosis and assemble the axial elements of the synaptonemal complex, however they

fail to complete homologous chromosome synapsis, resulting in meiotic germ cell apoptosis (Bourc'his and Bestor, 2004; Kuramochi-Miyagawa et al., 2004; Carmell et al., 2007; Öllinger et al., 2008). However, retrotransposable element specificity, and the timing of retrotransposon mis-expression seems to differ between *Tex19.1*$^{-/-}$, males and *Dnmt3L*$^{-/-}$, *Miwi2*$^{-/-}$ and *Mili*$^{-/-}$ males. Firstly, during the course of this project, I was able to show that retrotransposon upregulation at the RNA level appears to be specific to the *MMERVK10C* element in *Tex19.1*$^{-/-}$ testes (figure 3.4). This is based on expression analysis using repeat annotation of Illumina Mouse Whole Genome WG-6 Beadchip probes (repeat annotation and microarray analysis was performed by Ian Adams), which revealed an increase in abundance of *MMERVK10C* endogenous retrovirus transcripts but not other retrotransposons in *Tex19.1*$^{-/-}$ males. Genome wide expression analysis for retrotransposons has not been performed for *Dnmt3L*$^{-/-}$, *Miwi2*$^{-/-}$ and *Mili*$^{-/-}$ knockout males, however, it has been shown that all three mutants upregulate transcripts from the *IAP* endogenous retroviruses and *LINE1* elements, whereas *MMERVK10C* transcript levels do not change, at least in *Dnmt3L*$^{-/-}$ testes (Bourc'his and Bestor, 2004; Kuramochi-Miyagawa et al., 2004; Carmell et al., 2007; Öllinger et al., 2008; Michael Ling, personal communication). Furthermore, *MMERVK10C* transcripts are upregulated specifically in meiotic spermatocytes in *Tex19.1*$^{-/-}$ knockout mice but not in mitotic spermatogonia (Öllinger et al., 2008). In contrast, *IAP* and *LINE1* levels are upregulated in spermatogonia and spermatocytes in neonatal *Dnmt3L*$^{-/-}$, *Miwi2*$^{-/-}$ and *Mili*$^{-/-}$ testes supposedly due to impaired *de novo* methylation of *LINE1* and *IAP* elements in quiescent prospermatogonia (Bourc'his and Bestor, 2004; Kuramochi-Miyagawa et al., 2004; Carmell et al., 2007). Taken together, this data and previous analysis of the DNA methylation status of *MMERVK10C*, which did not identify any changes in *Tex19.1*$^{-/-}$ knockout testes, suggest that *Tex19.1* is not involved in the *Dnmt3L*/*Mili*/*Miwi2*-dependent mechanism that methylates and silences retrotransposons in quiescent prospermatogonia. In accordance with this, I have shown that *MMERVK10C* is not mis-expressed in *p53*$^{-/-}$*Dnmt1*$^{-/-}$ fibroblasts despite being hypomethylated which, together with Michael Ling's analysis of *MMERVK10C* expression in *Dnmt3L*$^{-/-}$ testes, implies that repression of *MMERVK10C* transcription in somatic cells and germ cells does not

rely as heavily on silencing by DNA methylation as *IAP* elements, for example (figure 3.5). These findings however, contrast with the observed increase of *MMERVK10C* expression in fibroblast cells treated with Aza, which is meant to induce DNA demethylation (figure 3.6). However, as described in chapter 3, it is possible that treatment with Aza induces further changes, in addition to DNA methylation, such as H3K9me^2 and possibly other histone marks (Wozniak et al., 2007). *MMERVK10C* silencing in ES cells was previously shown to be dependent on Kap1-mediated H3K9me^3, and given the fact that elements of the ERVK family are strongly enriched for this modification in fibroblasts, it is possible that *MMERVK10C* is also silenced by H3K9me^3 in somatic cells (Day et al., 2010). I was unable to confirm this hypothesis as I did not identify an upregulation *of MMERVK10C* in *Kap1*$^{-/-}$ pMEFs (figure 3.7). Analysis of *MMERVK10C* expression in fibroblasts with mutations in different histone methyltransferases would be useful to elucidate if this is indeed, one of the main mechanisms facilitating *MMEVRK10C* silencing. The biochemical function of *Tex19.1* is currently unknown. It appears unlikely that H3K9me^3 is globally affected in *Tex19.1*$^{-/-}$ testes as the observed RNA upregulation is largely restricted to *MMERVK10C.* It also seems unlikely that Tex19.1 only mediates H3K9me^3 at *MMERVK10C* loci. However, the chromatin state of *MMERVK10C* might change if the element is actively transcribed in the absence of *Tex19.1*. Changes in the chromatin state of *MMERVK10C* might result in the phenotype observed in *Tex19.1*$^{-/-}$ mutants. This is based on several mouse mutants which implicate histone modifications, such as H3K9me^2 and H3K9me^3, in the formation of homologous chromosome synapsis (Peters et al., 2001; Tachibana et al., 2007). For example, mice with mutations in the H3K9me^3 methyltransferases, Suv39h1 and Suv39h2, or the H3K9me2 methyltransferase, G9a, are all characterised by impaired homologous chromosome pairing and subsequently synapsis defects (Peters et al., 2001; Tachibana et al., 2007). It has been proposed by Takada et al. (2011) that H3K9me2 and H3K9me3 facilitate close alignment of centromeric regions of unpaired homologous chromosomes which in turn mediates progression of their pairing during early meiotic prophase. However, retrotransposon expression in these mutants has not been analysed and it would be very informative to perform qPCR for *MMERVK10C* and *IAP*

expression, both ERVK elements, in order to see if either element might be upregulated and a possible mediator or contributor to the observed phenotype. This is based on speculations by Öllinger et al. (2010) which propose that an altered chromatin state of retrotransposon elements might allow their participation in the homology search. Homology drives chromosome pairing and therefore participation of retrotransposons in the homology search, which are abundant and distributed throughout the genome, might mediate illegitimate recombination between non-homologous sequences (Öllinger et al., 2010; Romanish et al., 2010). It is possible that loss of transcriptional repression of *MMERVK10C* in *Tex19.1*$^{-/-}$ males results in a change of the chromatin environment of those elements with the consequence of illegitimate recombination between non-homologous sequences. One caveat with hypothesis is that early recombination foci are not depleted of retrotransposon sequences as one might expect if this theory should be true. As a matter of fact it has been reported that elements belonging to the MaLR retrotransposon superfamily are enriched at human and mouse recombination hot spots (Smagulova et al., 2011). Furthermore, in order to strengthen this hypothesis it would be important to determine if *MMERVK10C* transcription is indeed affected in *Tex19.1*$^{-/-}$ testes. Culture of *Tex19.1*$^{-/-}$ and control testes in presence of actinomycin D, an inhibitor of transcription, and subsequent qPCR analysis should give insights whether the increase of *MMERVK10C* transcript in *Tex19.1*$^{-/-}$ males is caused by an increase in transcription or RNA stability. However, upregulation of retrotransposon RNA might not only result from a defect in transcriptional inhibition but any stage of the retrotransposon life cycle. Inhibition lost at the DNA, RNA or protein level might all lead to increased retrotransposition events which in turn will result in the transcription of more elements and hence elevated RNA levels.

During the course of this thesis I also investigated if *Tex19.1* might act as a translational regulator of retrotransposons as increased stability of the *MMERVK10C* RNA by being more efficiently translated could explain the RNA upregulation observed in *Tex19.1*$^{-/-}$ testes. In accordance with this theory I was able to identify an upregulation of LINE1 ORF1p in *Tex19.1*$^{-/-}$ testes compared to controls but I was not able to observe a shift in

LINE1 or *MMERVK10C* RNA into the more translated part of a sucrose gradient (figures 3.8, 3.9 and 3.15). One of the shortcomings of the analysis presented here is that *MMERVK10C* protein expression could not be examined as functional antibodies are lacking. Sucrose gradient data and a poly A capture assay suggested that Tex19.1 is not associated with polysomes and does not bind to polyadenylated RNAs (figures 3.12 and 3.14). Furthermore, transfection of ES cells with reporter plasmids suggested that *LINE1* is not a target for translational inhibition by Tex19.1. Taken together those data, also when largely correlative, suggest that Tex19.1 is not a translational repressor of *MMERVK10C* or *LINE1* retrotransposons. Instead, it appears that Tex19.1 might repress retrotransposons at the level of protein stability. This hypothesis is based on the fact that Tex19.1 forms a stable complex with Ubr2 (figure 3.16) (Yang et al., 2010). In respect to the phenotypes of *Tex19.1*$^{-/-}$ mutant mice described in this thesis and by Öllinger et al. (2008) the interaction between Tex19.1 and Ubr2 was of great interest. The to date reported physiological functions of the N-end rule pathway (see also 3.1) are manifold and include the elimination of mis-folded proteins, signaling by transmembrane receptors, regulation of fat metabolism, function during apoptosis, chromosome segregation as well as bacterial and viral infections (Varshavsky, 2011). Tex19.1 and Ubr2 might target retrotransposons for degradation as it has been previously reported that the integrase protein of the human deficiency virus (HIV) represents an N-end rule substrate (Varshavsky, 2011). Other viral proteins, including retrotransposon proteins, might also be targets for N-end rule mediated proteolysis. Furthermore I have been able to show that Tex19.1 stimulates N-end rule mediated protein turnover when transfected into HeLa cells (figure 3.17). In order to directly test if LINE ORF1p stability is increased in the absence of *Tex19.1*, testes culture of *Tex19.1*$^{-/-}$ and control tissue in presence cycloheximide, an inhibitor of translation, and subsequent western blot analysis for LINE1 ORF1p might reveal if LINE1 ORF1p levels are elevated in the absence of Tex19.1 due to increased translation or indeed increased protein stability. Different mechanisms of retrotransposon suppression in male and female germ and ES cells might underlie the variation in the phenotypes between the sexes and cell types. *MMERVK10C* and other retrotransposons might be targets for N-recognins in male but not female germ

cells or alternatively other mechanisms might silence *MMERVK10C* in oocytes and therefore loss of N-end rule mediated repression would not result in a phenotype in the female germline. However, these are merely speculations and further work will be required to establish if these assumptions are true.

In conclusion the data I presented here suggest that Tex19.1 silences retrotransposons at the protein rather than the DNA or RNA level. I propose a hypothesis in which Tex19.1 mediates, together with Ubr2, the turnover of retroviral proteins. If this inhibition is lost increased retrotransposition and a subsequent increase in the numbers of transcribed elements might explain the increase in *MMERVK10C* abundance in *Tex19.1*$^{-/-}$ testes. Alternatively increased stability of retroviral proteins might result in an increase of retroviral particle formation which might in turn lead to increased stability of retroviral RNAs. The difference between the *LINE1* and *MMERVK10C* data, which is RNA upregulation of *MMERVK10C* but not *LINE1* RNA in *Tex19.1*$^{-/-}$ testes, might arise from the fact that other mechanisms might inhibit later stages of the *LINE1* but not *MMERVK10C* life cycle. It is currently not clear how many, if any, further retrotransposons are targets for Tex19.1 or in which way MMERVK10C or LINE1 are special so that they provide substrates for Tex19.1 but not for example IAP or IAPE. Increased retrotransposition or an increase in viral protein expression might both be able to cause defects in meiotic progression. Increased retrotransposition causes increased DNA damage. It has been shown very elegantly that *Mael*$^{-/-}$ mutants, which upregulate *IAP* and *LINE* retrotransposons and exhibit defects in chromosome synapsis, show extensive retrotransposition induced DNA damage, in absence of meiotic *Spo11*-dependent DSB by crossing *Spo11*$^{-/-}$ mutants with *Mael*$^{-/-}$ mice (Soper et al., 2008). DNA damage caused by retrotransposition events could interfere with normal meiotic recombination by attracting recombination proteins possibly resulting in attenuated levels of recombination proteins on *Spo11*-dependent DNA double strands breaks. Similarly, cytotoxicity of individual retrotransposon-encoded proteins may contribute to germ cell apoptosis in mouse mutants that correlate mis-expression of retrotransposons with defects in meiotic progression (Öllinger et al., 2010). However, it can not be

excluded that the relationship between upregulation of retrotransposons and aberrant progression through meiosis is indirect and an independent consequence of chromatin or gene expression changes. A retrotransposon knock-in model would provide a valuable tool to dissect the cause, consequence or indirect association between meiotic defects and retrotransposon mis-expression. In order to address this we aimed to generate a mouse model conditionally expressing *IAP* RNA. We chose *IAP* over other retrotransposons as it has been better characterized and more reagents are available to study this element compared to other retrotransposons in the mouse genome. I generated an *IAP* expressing construct in a pROSA targeting vector. Transient transfection into ES cells showed that the construct is functional. Unfortunately, we have been unsuccessful to generate correctly targeted ES cell clones so far. However, a mouse model with ectopic retrotransposon expression will provide an important tool in order to understand the relationship between retrotransposon upregulation and chromosome segregation and therefore this project should be further pursued.

6.2 Aneuploidy in *Tex19.1*$^{-/-}$ ES cells and decreased viability of *Tex19.1*$^{-/-}$ embryos

6.2.1 Aneuploidy in *Tex19.1*$^{-/-}$ ES cells

In addition to its previous attributed role in promoting genetic and chromosomal stability during spermatogenesis I provide preliminary evidence that *Tex19.1* might also be required to facilitate faithful chromosome segregation in pluripotent cells (figure 4.7). Interestingly the interaction partner of Tex19.1, Ubr2 has previously been implicated in faithful chromosome segregation during mitosis. Although the underlying mechanisms remain unclear, it has been shown that *Ubr2*$^{-/-}$ fibroblasts display genome instability caused by spontaneous gaps or breaks in metaphase chromosomes as well as an increase of chromosome mis-segregation at anaphase (Ouyang et al., 2006). The factors that might be involved in mediating the role of *Ubr2* in promoting genetic and chromosomal stability include C-terminal fragments of the Scc1p/Rad21 cohesin subunit (Varshavsky,

2011). In eukaryotes C-terminal fragments of the Scc1p/Rad21 cohesin subunit are produced by separase through cleavage of Scc1p/Rad21 at the metaphase to anaphase transition to allow sister chromatid separation (figure 6.1) (Uhlmann et al., 1999; Uhlmann et al., 2000; Rao et al., 2001). Scc1p/Rad21 C-terminal fragments are normally short lived (t1/2<2 min) due to degradation by ubiquitin/proteasome dependent N-end rule pathway (Rao et al., 2001). In *ubr1Δ S. cerevisiae* cells, which lack the N-end rule pathway as *ubr1*, a homologue of Ubr1 and Ubr2, represents the sole N-recognin in yeast, the Scc1p fragment appears to be metabolically stabilized resulting in increased aneuploidy and cell lethality (Rao et al., 2001). Similarly, overexpression of Scc1p fragments with mutations in the destabilizing residue, which generates a long-lived derivative of the Scc1p fragment, causes increased frequency of chromosome loss and compromised cell viability (Rao et al., 2001). Those data suggest that removal of this cohesin fragment is critical to ensure high fidelity of chromosome cohesion and segregation. It has been postulated that rapid degradation of the cleaved cohesin protein ensures that only uncleaved Scc1p/Rad21 proteins are incorporated into reconstituting cohesin complexes (Rao et al., 2001; Varshavsky, 2011). It is plausible to speculate that at least part of the *Tex19.1*$^{-/-}$ phenotypes might be mediated by defects in the N-end rule pathway and future work should explore this possibility. This is based on the observed interaction between *Ubr2* and *Tex19.1* and the similarities in the phenotypes of mutants of either gene which suggest that Tex19.1 might mediate Ubr2 dependent N-end rule photolytic degradation of target proteins (figure 3.17) (Yang et al., 2010). This is supported by the fact that *Tex19.1* appears to promote N-end rule dependent protein degradation in HeLa cells (figure 3.18). Accordingly defects in mitotic chromosome segregation in *Tex19.1*$^{-/-}$ pluripotent cells might be mediated by defects in N-end rule protein degradation in the absence of Tex19.1. In order to show that this mechanism is responsible for the aneuploidy in *Tex19.1*$^{-/-}$ ES cells, *Tex19.1*$^{-/-}$ and control ES cells should be analysed and compared for protein levels of the Rad21 C-terminal fragment. Overexpression of a long-lived Rad21 C-terminal fragment in ES cells followed by karyotyping should reveal if rapid turnover of cleaved cohesin is required for fidelity of

chromosome segregation in pluripotent cells. The above described data obtained from *S. cerevisiae* suggest that this might be the case.

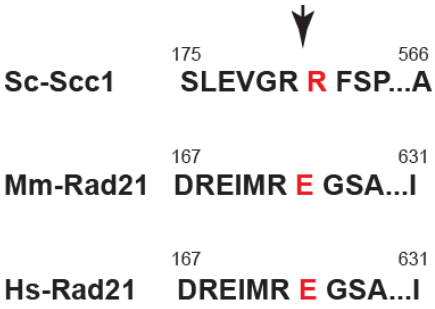


Figure 6.1: Alignment of separase target sites and by cleavage generated N-end rule degradation signals for Scc1. Amino acid residues shown in black as single-letter abbreviations. Arrowhead and red d residues indicate the cleavage sites and N-terminal residues of the corresponding C-terminal fragments. Sc, Mm and Hs refer to *S. cerevisiae*, *M. musculus* and *H. sapiens*, respectively. R is a primary de-stabilising residue, E a secondary de-stabilising residue. Figure was taken and modified from Varshavsky (2011).

6.2.2 Decreased viability of *Tex19.1*$^{-/-}$ embryos

The data I present in chapter 4 also demonstrate that absence of Tex19.1 compromises viability of embryos of both sexes (table 4.1). I show that lethality of *Tex19.1*$^{-/-}$ males and females is exacerbated in pups born from lactating mothers but with females being considerably stronger affected than males (table 4.3). It is currently not clear which mechanisms cause the here observed problems. One of the weaknesses of this study is that I am not able to distinguish if embryonic lethality is increased in response to embryonic diapause or in response to lactation, as the analysis of chemically induced diapaused embryos was inconclusive. Furthermore, chemically induced diapause might differ from lactation induced diapause as I observed a strong reduction of all females irrespective of genotype in chemically diapaused embryos which is not observed in naturally diapaused embryos. However, no comparable studies about this matter exist to date. Similarly to *Tex19.1*$^{-/-}$ mice, *Ubr2*$^{-/-}$ mice exhibit male infertility, female subfertility and embryonic lethality of mutant females (Kwon et al., 2003; Öllinger et al., 2008; An et al., 2009; Yang et al., 2010; chapters 4 and 5 of this thesis). In order to test if

aneuploidy is increased in *Tex19.1*$^{-/-}$ embryos, as observed in *Tex19.1*$^{-/-}$ ES cells, karyotyping of *Tex19.1*$^{-/-}$ and control diapaused blastocysts needs to be performed. Karyotyping of blastocysts will be technically challenging due to such a small tissue source. However, karyotyping of blastocyst cells has been previously reported (Lightfoot et al., 2005). DNA FISH analysis using a *Neo* probe in order to help to distinguish between *Tex19.1*$^{+/+}$, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ embryos, as demonstrated by Spencer and colleagues (2011), could be used for blastocyst genotyping. DNA FISH for major satellites could be applied at the same time in order to facilitate chromosome counts (see chapter 5). Sectioning of early post-implantation embryos and terminal deoxynucleotidyl transferase end labelling (TUNEL) in order to detect DNA fragmentation, a well known biomarker for apoptosis in tissue, should be performed for *Tex19.1*$^{-/-}$ and control embryos in order to analyse if there is indeed an increase in apoptosis in *Tex19.1*$^{-/-}$ mutants post-implantation as would be expected for aneuploid embryos (Gavrieli et al., 1992; Lightfoot et al., 2006). However, it is not clear if the aneuploidy observed in *Tex19.1*$^{-/-}$ ES cells is representative for the role of Tex19.1 during embryonic development. Based on the observation that arrested blastocysts contain fewer cycling cells compared to nondelayed embryos it appears unlikely that *Tex19.1*$^{-/-}$ embryos accumulate high rates of aneuploidy during diapause (Given, 1988; Given and Weitlauf, 1981; Battle-Morera et al., 2008). Furthermore, it is unclear how a chromosome segregation defect would affect female embryos stronger during diapause than male embryos as observed for *Tex19.1*$^{-/-}$ embryos.

Alternatively, based on observations from the male germline, *Tex19.1* might function in suppression of retrotransposons in pluripotent cells during embryonic development. The male-female sex bias might be caused by arrested female *Tex19.1*$^{-/-}$ embryos being more vulnerable to loss of post-transcriptional inhibition of retrotransposons by absence of *Tex19.1* than *Tex19.1*$^{-/-}$ males. This could result from a potent retrotransposon on the X chromosome or increased retrotransposon expression in XX pluripotent cells as XX ES cells have been shown to be hypomethylated compared to XY ES cells (Zvetkova et al., 2005). However, there is no evidence for this hypothesis at the moment.

Causes known to affect the viability of female embryos involve defects in X-chromosome inactivation. X-chromosome inactivation (XCI) achieves balance of the X-chromosome dosage between XX and XY individuals during embryogenesis (Augui et al., 2011). In mice an imprinted process is initiated during the early cleavage stages which results in the inactivation of the paternal X-chromosome (Xp). Xp silencing is maintained in extraembryonic tissues but reversed in the inner cells mass where a subsequent wave of a random form of XCI occurs which silences either the Xp or the maternal X-chromosome (Xm) (Okamoto and Heard, 2006). Mutations that affect dosage compensation result in female lethality during early post-implantation development or before 13.5dpc depending on the defect (Okamoto and Heard, 2006; Augui et al., 2011). In addition to the N-end rule pathway Ubr2 is also involved in monoubiquitination of histone 2A (H2A) (An et al., 2010). H2A monoubiquitination mediated by Ring1B is known to be required for XCI (de Napoles et al., 2004). It is possible that Tex19.1 and Ubr2 also play a role in XCI. However, it is not clear why defects in dosage compensation should be exacerbated during embryonic diapause. Furthermore, increased lethality of *Tex19.1*$^{-/-}$ males during diapause suggests that mechanisms other than XCI at least contribute to the phenotype described in chapter 4 (table 4.3).

6.3 Increased levels of aneuploidy in the germline of *Tex19.1*$^{-/-}$ females

During the course of this project I was able to demonstrate that *Tex19.1* not only promotes genetic and chromosomal stability during spermatogenesis, as previously reported, and possibly early embryonic development, but also during female meiosis. The data presented here suggest that in absence of *Tex19.1* sister chromatid cohesion in mouse oocytes might be prematurely lost resulting in chromosome mis-segregation and aneuploid offspring. As described under 1.6, premature sister chromatid separation as well as non-disjunction of whole bivalents at meiosis I have been postulated to

contribute to the generation of human aneuploidy. Which of the two is the more prevalent mechanism is still subject to debate (see 1.6). Analysis and karyotyping of meiosis I human oocytes has been extremely limited due to the fact that meiosis I oocytes are difficult to obtain and most human oocytes are collected at the metaphase II stage after super ovulation. Studies of human trisomies have provided valuable insights regarding the mechanisms that result in human aneuploidy. However, analysis of trisomies is not able to distinguish if the abnormal karyotype has been generated through a non-disjunction event or premature sister chromatid separation (figures 1.14 and 1.15). Premature sister chromatid separation observed in human metaphase II oocytes might be caused by a loss of sister chromatid cohesin or be the result of univalents achieving bi-orientation while aligning to the metaphase plate, which could lead to the equational segregation of sister chromatids to opposite poles (Kouznetsova et al., 2007). This demonstrates that the relationship between non-disjunction, premature sister chromatid separation and the occurrence of aneuploidy might be difficult to dissect. The data presented in this thesis suggest that *Tex19.1*$^{-/-}$ females are aneuploid due to premature separation of sister chromatids (figure 5.9). If this hypothesis can be confirmed, *Tex19.1*$^{-/-}$ oocytes would provide a valuable tool to dissect the cause and consequences of premature sister chromatid separation during meiosis. Several mouse mutants for SAC genes and nondisjunction have been described previously (Edelmann et al., 1996; Yuan et al., 2002; Revenkova et al., 2004; Hodges et al., 2005; Homer et al., 2005). A few mouse mutants attribute aneuploidy to premature sister chromatid separation (Revenkova et al., 2004; Leland et al., 2009). The *Tex19.1*$^{-/-}$ females might provide a valuable model for the study of the mechanisms that lead to premature sister chromatid separation and therewith advance our understanding of the causes of human aneuploidy. The biggest limitation of the study presented here is that the aneuploidy data solely rely on chromosome counts from fixed oocytes. I found that artificial chromosome loss, chromosome superposition and poor morphology were common issues with this method. The high rates of hypoploidy observed in both *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ oocytes are probably attributable to an artificial loss of chromosomes during the fixation procedure. However, in order to strengthen the observations made in chapter 5 several chromosome

spreads were counted independently, without knowledge of the corresponding genotype, by myself and Ian Adams. Furthermore, consistent aneuploidy levels around 30%, which are sufficient to explain the reduction in litter size observed, in 3 different stages of oocyte development suggest that *Tex19.1*$^{-/-}$ oocytes indeed exhibit defects in meiotic chromosome segregation. Further work should concentrate on elucidating the molecular mechanisms that underlie the chromosome segregation defects observed in *Tex19.1*$^{-/-}$ female germ cells. As described under 1.4 correct meiotic chromosome behavior requires the replacement of mitotic cohesin proteins by meiotic cohesins. For example Rad21 is largely replaced by Rec8. However, IF studies of meiotic male germ cells have shown that the mitotic cohesin proteins Stag2 and Rad21 are localised to meiotic chromosomes during the early stages of prophase I and are still present at the dipolotene stage (Prieto et al., 2002). It is possible that delayed turnover of Rad21 C-terminal fragments leads to incorporation of this unfunctional protein fragment into the meiotic cohesin complexes resulting in defects in sister chromatid cohesin as observed in the *Tex19.1*$^{-/-}$ female germline. Alternatively Rad21 C-terminal fragments might interfere with cohesin loading onto the chromosomes. Another possibility is that Rec8 C-terminal fragments are targets for Tex19.1 and Ubr2. Rec8 C-terminal fragments produced by seperase cleavage during the metaphase I to anaphase I transition might associate with the remaining cohesin complexes that facilitate centromeric cohesion. This might replace some of the bound cohesin proteins and weaken sister chromatid cohesin at the centromeres which might result in premature sister chromatid separation. Another possibility is that Rec8 fragments after seperase cleavage associate with *Sgol2* which might interfere with the protection of centromeric cohesin from cleavage. One of the caveats with this theory is that the Rec8 C-terminal fragments do not contain an N-degron but, according to the N-end rule, stabilising residues (Kudo et al., 2009). The cleavage sites for Rec8 have been determined by *in vitro* analysis. Two of the C-terminal fragments produced by separase cleavage of Rec8 have alanine as their N-terminal residue which is a stabilising amino acid. Only the smallest one has glutamic acid as N-terminal fragment which is a secondary de-stabilising residue. It is possible that Tex19.1 might be involved in targeting Ubr2 to proteins that do not contain N-terminal residues,

such as mouse or human Rec8, during meiosis. If Rec8 should indeed represent a target for Tex19.1 and Ubr2, this could also explain why *S. cerevisiae* can cope without Tex19.1.This is based on the fact that the two cleavage products of Rec8 after seperase cleavage in yeast have lysine as N-terminal residue which is a primary de-stabilising amino acid. It would be interesting to examine the N-terminal residues of cohesin fragments of further non-mammalian species, which do not posses Tex19.1, to see whether or not those fragments have N-degrons. To test if N-end rule mediated protein degradation is affected in *Tex19.1*$^{-/-}$ oocytes, mRNA produced from the GFP-reporters described in chapter 1 should be injected into *Tex19.1*$^{-/-}$ and control oocytes. In addition injection of mRNA expressing Rad21 or Rec8 long-lived C-terminal fragments into wild type oocytes would help to elucidate if stability of the mitotic or meiotic cohesin fragments can cause defects in chromosome segregation during meiosis.

The main aim of my PhD project was the phenotypic characterization of *Tex19.1*$^{-/-}$ females. I have been able to show that *Tex19.1*$^{-/-}$ females are subfertile and transmit high rates of aneuploidy to their offspring. This increase in aneuploidy in *Tex19.1*$^{-/-}$ oocytes presumably results from a defect during the first meiotic division and might be caused by premature separation of sister chromatids. I propose a hypothesis which suggests that this might be caused by insufficient turnover of non-functional cohesin protein fragments which are incorporated into the meiotic cohesin complex. This model is summarized in figure 6.2. Weakened sister chromatid cohesin could then cause premature separation during the first meiotic division resulting in the formation of aneuploid oocytes.

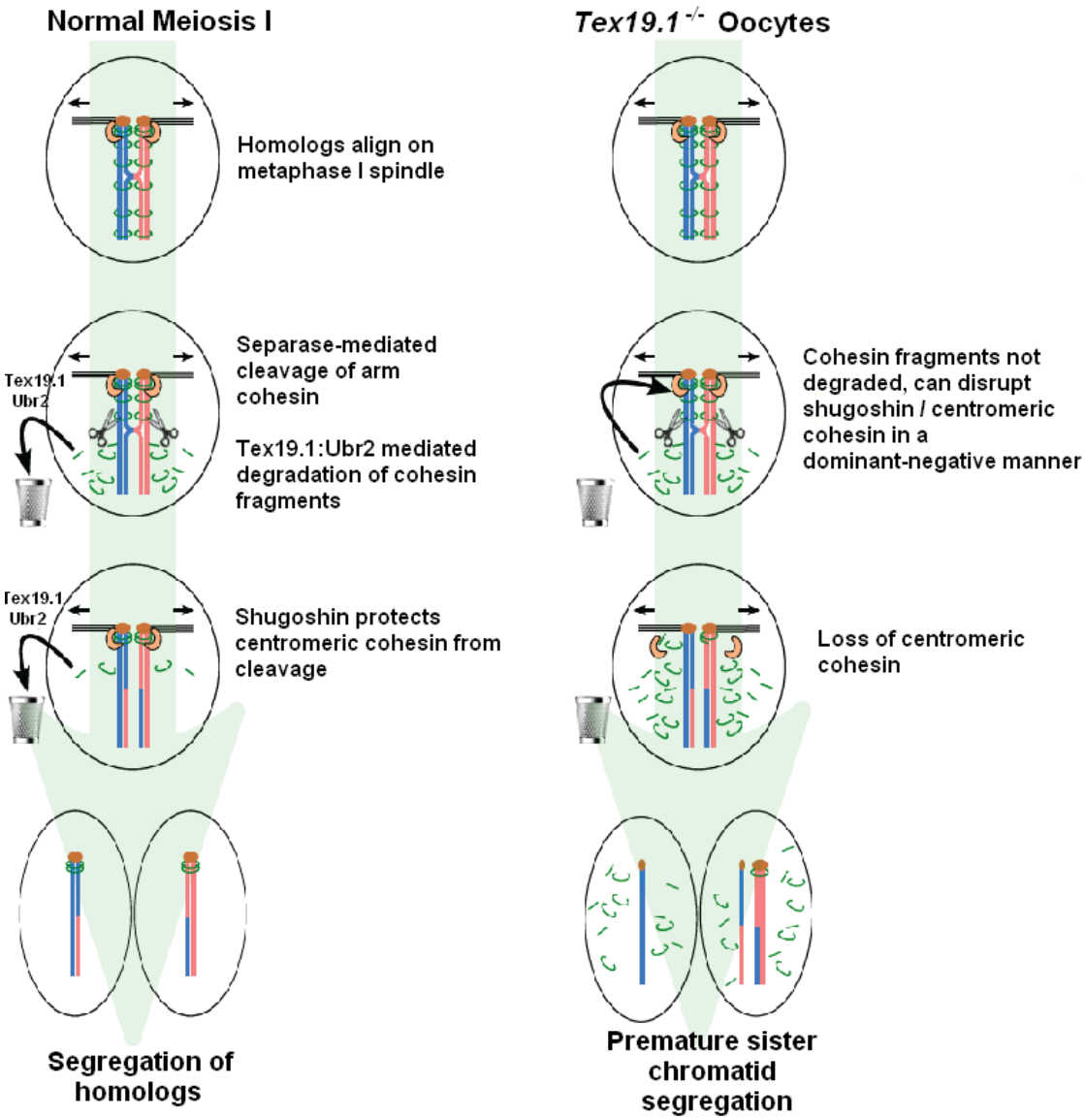


Figure 6.2 Proposed model for the role of Tex19.1 during female meiosis (adapted from Ian Adams).

6.4 Conclusion

In conclusion, the data presented in this thesis suggest that *Tex19.1* promotes, in addition to its previous attributed role during spermatogenesis, genetic and chromosomal stability during female germline development, and possibly during early embryonic development. The next step will be to elucidate the molecular/biochemical function of *Tex19.1* in order to fully appreciate the *Tex19.1*$^{-/-}$ phenotypes characterised during this project. A role for Tex19.1 in the N-end rule pathway would suggest that certain proteins might be stabilised in the absence of Tex19.1. Increased stability of some of these proteins, such as cohesin and/or retrotransposon protein fragments, could potentially mediate, or modulate, the observed *Tex19.1*$^{-/-}$ phenotypes. It is possible that the embryonic lethality as well as the male and female fertility phenotypes are caused by defects in N-end rule mediated protein turn over. Given the high rate of aneuploidy observed among early pregnancies in humans understanding of the mechanisms that influence chromosome segregation during gametogensis and early embryonic development is of critical importance to provide information regarding the occurrence of aneuploidy and the resulting reproductive diseases in our species. The *Tex19.1*$^{-/-}$ mice might provide a valuable model system to advance our understanding of chromosome mis-segregation during mitosis and meiosis and possibly mechanisms that underlie the high incidence of aneuploidy in human oocytes.

References

Adams, I. R. and A. McLaren (2002). "Sexually dimorphic development of mouse primordial germ cells: switching from oogenesis to spermatogenesis." Development **129**(5): 1155-64.

Aittomaki, K., J. L. Lucena, et al. (1995). "Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure." Cell **82**(6): 959-68.

Akande, V., C. Turner, et al. (2009). "Impact of Chlamydia trachomatis in the reproductive setting: British Fertility Society Guidelines for practice." Hum Fertil (Camb) **13**(3): 115-25.

Alarcon, V. B. and Y. Marikawa (2003). "Deviation of the blastocyst axis from the first cleavage plane does not affect the quality of mouse postimplantation development." Biol Reprod **69**(4): 1208-12.

Alarcon, V. B. and Y. Marikawa (2005). "Unbiased contribution of the first two blastomeres to mouse blastocyst development." Mol Reprod Dev **72**(3): 354-61.

Albertini, D. F. and S. L. Barrett (2003). "Oocyte-somatic cell communication." Reprod Suppl **61**: 49-54.

Ambartsumyan, G. and A. T. Clark (2008). "Aneuploidy and early human embryo development." Hum Mol Genet **17**(R1): R10-5.

An, J. Y., E. A. Kim, et al. (2010). "UBR2 mediates transcriptional silencing during spermatogenesis via histone ubiquitination." Proc Natl Acad Sci U S A **107**(5): 1912-7.

An, J. Y., J. W. Seo, et al. (2006). "Impaired neurogenesis and cardiovascular development in mice lacking the E3 ubiquitin ligases UBR1 and UBR2 of the N-end rule pathway." Proc Natl Acad Sci U S A **103**(16): 6212-7.

Angell, R. (1997). "First-meiotic-division nondisjunction in human oocytes." Am J Hum Genet **61**(1): 23-32.

Angell, R. R. (1991a). "Predivision in human oocytes at meiosis I: a mechanism for trisomy formation in man." Hum Genet **86**(4): 383-7.

Angell, R. R., W. Ledger, et al. (1991b). "Cytogenetic analysis of unfertilized human oocytes." Hum Reprod **6**(4): 568-73.

Angell, R. R., J. Xian, et al. (1994). "First meiotic division abnormalities in human oocytes: mechanism of trisomy formation." Cytogenet Cell Genet **65**(3): 194-202.

Anson-Cartwright, L., K. Dawson, et al. (2000). "The glial cells missing-1 protein is essential for branching morphogenesis in the chorioallantoic placenta." Nat Genet **25**(3): 311-4.

Antonarakis, S. E., R. Lyle, et al. (2004). "Chromosome 21 and down syndrome: from genomics to pathophysiology." Nat Rev Genet **5**(10): 725-38.

Aravin, A., D. Gaidatzis, et al. (2006). "A novel class of small RNAs bind to MILI protein in mouse testes." Nature **442**(7099): 203-7.

Augui, S., E. P. Nora, et al. (2011). "Regulation of X-chromosome inactivation by the X-inactivation centre." Nat Rev Genet **12**(6): 429-42.

Aziz, M. and H. Alexandre (1991). "The origin of the nascent blastocoele in preimplantation mouse embryos ultrastructural cytochemistry and effect of chloroquine." Development Genes and Evolution **200**(2): 77-85

Baudat, F., J. Buard, et al. (2010). "PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice." Science **327**(5967): 836-40.

Baudat, F., K. Manova, et al. (2000). "Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking Spo11." Mol Cell **6**(5): 989-98.

Bedford, L., S. Paine, et al. (2010). "Assembly, structure, and function of the 26S proteasome." Trends Cell Biol **20**(7): 391-401.

Beechey, C. V. and A. G. Searle (1988). "Effects of zero to four copies of chromosome 15 on mouse embryonic development." Cytogenet Cell Genet **47**(1-2): 66-71.

Bellve, A. R., J. C. Cavicchia, et al. (1977). "Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization." J Cell Biol **74**(1): 68-85.

Bendel-Stenzel, M., R. Anderson, et al. (1998). "The origin and migration of primordial germ cells in the mouse." Semin Cell Dev Biol **9**(4): 393-400.

Ben-Haim, N., C. Lu, et al. (2006). "The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4." Dev Cell **11**(3): 313-23.
Bernstein, K. A. and R. Rothstein (2009). "At loose ends: resecting a double-strand break." Cell **137**(5): 807-10.
Bischoff, M., D. E. Parfitt, et al. (2008). "Formation of the embryonic-abembryonic axis of the mouse blastocyst: relationships between orientation of early cleavage divisions and pattern of symmetric/asymmetric divisions." Development **135**(5): 953-62.
Bishop, C. E., D. J. Whitworth, et al. (2000). "A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse." Nat Genet **26**(4): 490-4.
Boheler, K. R., J. Czyz, et al. (2002). "Differentiation of pluripotent embryonic stem cells into cardiomyocytes." Circ Res **91**(3): 189-201.
Bolcun-Filas, E., Y. Costa, et al. (2007). "SYCE2 is required for synaptonemal complex assembly, double strand break repair, and homologous recombination." J Cell Biol **176**(6): 741-7.
Bolcun-Filas, E., E. Hall, et al. (2009). "Mutation of the mouse Syce1 gene disrupts synapsis and suggests a link between synaptonemal complex structural components and DNA repair." PLoS Genet **5**(2): e1000393.
Borts, R. H., S. R. Chambers, et al. (2000). "The many faces of mismatch repair in meiosis." Mutat Res **451**(1-2): 129-50.
Bourc'his, D. and T. H. Bestor (2004). "Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L." Nature **431**(7004): 96-9.
Boyle, S., S. Gilchrist, et al. (2001). "The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells." Hum Mol Genet **10**(3): 211-9.
Braude, P., V. Bolton, et al. (1988). "Human gene expression first occurs between the four- and eight-cell stages of preimplantation development." Nature **332**(6163): 459-61.
Broman, K. W., L. B. Rowe, et al. (2002). "Crossover interference in the mouse." Genetics **160**(3): 1123-31.
Brunmeir, R., S. Lagger, et al. (2010). "Epigenetic regulation of a murine retrotransposon by a dual histone modification mark." PLoS Genet **6**(4): e1000927.
Bugge, M., A. Collins, et al. (2007). "Non-disjunction of chromosome 13." Hum Mol Genet **16**(16): 2004-10.
Bugge, M., A. Collins, et al. (1998). "Non-disjunction of chromosome 18." Hum Mol Genet **7**(4): 661-9.
Burgoyne, P. S., S. K. Mahadevaiah, et al. (2007). "The management of DNA double-strand breaks in mitotic G2, and in mammalian meiosis viewed from a mitotic G2 perspective." Bioessays **29**(10): 974-86.
Burgoyne, P. S., S. K. Mahadevaiah, et al. (2009). "The consequences of asynapsis for mammalian meiosis." Nat Rev Genet **10**(3): 207-16.
Cacheiro, N. L., J. C. Rutledge, et al. (1994). "Cytogenetic analysis of malformed mouse fetuses derived from balanced translocation heterozygotes." Cytogenet Cell Genet **66**(2): 139-48.
Calabrese, J. M., A. C. Seila, et al. (2007). "RNA sequence analysis defines Dicer's role in mouse embryonic stem cells." Proc Natl Acad Sci U S A **104**(46): 18097-102.
Cameron, F. J. and A. H. Sinclair (1997). "Mutations in SRY and SOX9: testis-determining genes." Hum Mutat **9**(5): 388-95.
Capo-Chichi, C. D., M. E. Rula, et al. (2005). "Perception of differentiation cues by GATA factors in primitive endoderm lineage determination of mouse embryonic stem cells." Dev Biol **286**(2): 574-86.
Carmell, M. A., A. Girard, et al. (2007). "MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline." Dev Cell **12**(4): 503-14.
Carpenter, A. T. (1994). "Chiasma function." Cell **77**(7): 957-62.
Carrel, L. and H. F. Willard (2005). "X-inactivation profile reveals extensive variability in X-linked gene expression in females." Nature **434**(7031): 400-4.
Carrell, D. T., L. Liu, et al. (2005). "Comparison of maturation, meiotic competence, and chromosome aneuploidy of oocytes derived from two protocols for in vitro culture of mouse secondary follicles." J Assist Reprod Genet **22**(9-10): 347-54.

Castaneda, J., P. Genzor, et al. (2011). "piRNAs, transposon silencing, and germline genome integrity." Mutat Res.
Chambers, I., D. Colby, et al. (2003). "Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells." Cell **113**(5): 643-55.
Chassot, A. A., F. Ranc, et al. (2008). "Activation of beta-catenin signaling by Rspo1 controls differentiation of the mammalian ovary." Hum Mol Genet **17**(9): 1264-77.
Chawengsaksophak, K., W. de Graaff, et al. (2004). "Cdx2 is essential for axial elongation in mouse development." Proc Natl Acad Sci U S A **101**(20): 7641-5.
Chazaud, C., Y. Yamanaka, et al. (2006). "Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway." Dev Cell **10**(5): 615-24.
Chen, L., A. Yabuuchi, et al. (2009). "Cross-regulation of the Nanog and Cdx2 promoters." Cell Res **19**(9): 1052-61.
Chen, R. Z., U. Pettersson, et al. (1998). "DNA hypomethylation leads to elevated mutation rates." Nature **395**(6697): 89-93.
Chen, T., Y. Ueda, et al. (2003). "Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b." Mol Cell Biol **23**(16): 5594-605.
Cheng, E. Y., P. A. Hunt, et al. (2009). "Meiotic recombination in human oocytes." PLoS Genet **5**(9): e1000661.
Chong, S. and E. Whitelaw (2004). "Epigenetic germline inheritance." Curr Opin Genet Dev **14**(6): 692-6.
Chow, J. C., Z. Yen, et al. (2005). "Silencing of the mammalian X chromosome." Annu Rev Genomics Hum Genet **6**: 69-92.
Chroscicka, A., S. Komorowski, et al. (2004). "Both blastomeres of the mouse 2-cell embryo contribute to the embryonic portion of the blastocyst." Mol Reprod Dev **68**(3): 308-12.
Chuma, S., M. Hosokawa, et al. (2006). "Tdrd1/Mtr-1, a tudor-related gene, is essential for male germ-cell differentiation and nuage/germinal granule formation in mice." Proc Natl Acad Sci U S A **103**(43): 15894-9.
Collins, I. and C. S. Newlon (1994). "Meiosis-specific formation of joint DNA molecules containing sequences from homologous chromosomes." Cell **76**(1): 65-75.
Constantinou, A., A. A. Davies, et al. (2001). "Branch migration and Holliday junction resolution catalyzed by activities from mammalian cells." Cell **104**(2): 259-68.
Costa, Y. and H. J. Cooke (2007). "Dissecting the mammalian synaptonemal complex using targeted mutations." Chromosome Res **15**(5): 579-89.
Costa, Y., R. Speed, et al. (2005). "Two novel proteins recruited by synaptonemal complex protein 1 (SYCP1) are at the centre of meiosis." J Cell Sci **118**(Pt 12): 2755-62.
Cross, J. C., D. G. Simmons, et al. (2003). "Chorioallantoic morphogenesis and formation of the placental villous tree." Ann N Y Acad Sci **995**: 84-93.
Dalcq (1957). Introduction to General Embryology, Oxford Univ. Press, London.
Dantuma, N. P., K. Lindsten, et al. (2000). "Short-lived green fluorescent proteins for quantifying ubiquitin/proteasome-dependent proteolysis in living cells." Nat Biotechnol **18**(5): 538-43.
Davis, C. M., P. G. Constantinides, et al. (1989). "Activation and demethylation of the intracisternal A particle genes by 5-azacytidine." Cell Differ Dev **27**(2): 83-93.
Day, D. S., L. J. Luquette, et al. (2010). "Estimating enrichment of repetitive elements from high-throughput sequence data." Genome Biol **11**(6): R69.
de Cuevas, M., J. K. Lee, et al. (1996). "alpha-spectrin is required for germline cell division and differentiation in the Drosophila ovary." Development **122**(12): 3959-68.
De La Fuente, R., C. Baumann, et al. (2006). "Lsh is required for meiotic chromosome synapsis and retrotransposon silencing in female germ cells." Nat Cell Biol **8**(12): 1448-54.
de Napoles, M., J. E. Mermoud, et al. (2004). "Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation." Dev Cell **7**(5): 663-76.
de Napoles, M., T. Nesterova, et al. (2007). "Early loss of Xist RNA expression and inactive X chromosome associated chromatin modification in developing primordial germ cells." PLoS One **2**(9): e860.
de Rooij, D. G. and J. A. Grootegoed (1998). "Spermatogonial stem cells." Curr Opin Cell Biol **10**(6): 694-701.

de Rooij, D. G. and S. C. Mizrak (2008). "Deriving multipotent stem cells from mouse spermatogonial stem cells: a new tool for developmental and clinical research." Development **135**(13): 2207-13.
de Rooij, D. G. and L. D. Russell (2000). "All you wanted to know about spermatogonia but were afraid to ask." J Androl **21**(6): 776-98.
de Ruijter, A. J., A. H. van Gennip, et al. (2003). "Histone deacetylases (HDACs): characterization of the classical HDAC family." Biochem J **370**(Pt 3): 737-49.
de Vries, F. A., E. de Boer, et al. (2005). "Mouse Sycp1 functions in synaptonemal complex assembly, meiotic recombination, and XY body formation." Genes Dev **19**(11): 1376-89.
DesGroseilliers, M., E. Lemyre, et al. (2002). "Tetrasomy Y by structural rearrangement: clinical report." Am J Med Genet **111**(4): 401-4.
Di Giacomo, M., M. Barchi, et al. (2005). "Distinct DNA-damage-dependent and -independent responses drive the loss of oocytes in recombination-defective mouse mutants." Proc Natl Acad Sci U S A **102**(3): 737-42.
Dietrich, J. E. and T. Hiiragi (2007). "Stochastic patterning in the mouse pre-implantation embryo." Development **134**(23): 4219-31.
Donahue, R. P. (1968). "Maturation of the mouse oocyte in vitro. I. Sequence and timing of nuclear progression." J Exp Zool **169**(2): 237-49.
Dong, J., D. F. Albertini, et al. (1996). "Growth differentiation factor-9 is required during early ovarian folliculogenesis." Nature **383**(6600): 531-5.
Dong, K. B., I. A. Maksakova, et al. (2008). "DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity." Embo J **27**(20): 2691-701.
Dyban, A. P. a. B., V.S. (1987). Cytogenetics of mammalian embryonic development. Oxford and Oxford and New York, Clarendon Press
Edelmann, W., P. E. Cohen, et al. (1996). "Meiotic pachytene arrest in MLH1-deficient mice." Cell **85**(7): 1125-34.
Edelmann, W., P. E. Cohen, et al. (1999). "Mammalian MutS homologue 5 is required for chromosome pairing in meiosis." Nat Genet **21**(1): 123-7.
Edson, M. A., A. K. Nagaraja, et al. (2009). "The mammalian ovary from genesis to revelation." Endocr Rev **30**(6): 624-712.
Eichenlaub-Ritter, U. (2005). "Mouse genetic models for aneuploidy induction in germ cells." Cytogenet Genome Res **111**(3-4): 392-400.
Eijpe, M., C. Heyting, et al. (2000). "Association of mammalian SMC1 and SMC3 proteins with meiotic chromosomes and synaptonemal complexes." J Cell Sci **113 (Pt 4)**: 673-82.
Eijpe, M., H. Offenberg, et al. (2003). "Meiotic cohesin REC8 marks the axial elements of rat synaptonemal complexes before cohesins SMC1beta and SMC3." J Cell Biol **160**(5): 657-70.
Epstein, C. J. and B. Travis (1979). "Preimplantation lethality of monosomy for mouse chromosome 19." Nature **280**(5718): 144-5.
Eskeland, R., B. Czermin, et al. (2004). "The N-terminus of Drosophila SU(VAR)3-9 mediates dimerization and regulates its methyltransferase activity." Biochemistry **43**(12): 3740-9.
Esposito, A. M., M. Mateyak, et al. (2010). "Eukaryotic polyribosome profile analysis." J Vis Exp(40).
Fenwick, M. L. (1968). "The effect of ribonuclease on polysomes and ribosomes of bacteria and animal cells." Biochem J **107**(4): 481-9.
Filis, P., T. Lannagan, et al. (2009). "Phospholipase C-beta1 signaling affects reproductive behavior, ovulation, and implantation." Endocrinology **150**(7): 3259-66.
Flach, G., M. H. Johnson, et al. (1982). "The transition from maternal to embryonic control in the 2-cell mouse embryo." Embo J **1**(6): 681-6.
Fleming, T. P., B. Sheth, et al. (2001). "Cell adhesion in the preimplantation mammalian embryo and its role in trophectoderm differentiation and blastocyst morphogenesis." Front Biosci **6**: D1000-7.
Ford, C. E., E. P. Evans, et al. (1975). "A functional 'sex-reversed' oocyte in the mouse." Proc R Soc Lond B Biol Sci **190**(1099): 187-97.
Froenicke, L., L. K. Anderson, et al. (2002). "Male mouse recombination maps for each autosome identified by chromosome painting." Am J Hum Genet **71**(6): 1353-68.
Frost, R. J., F. K. Hamra, et al. (2010). "MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs." Proc Natl Acad Sci U S A **107**(26): 11847-52.

Fujimori, T., Y. Kurotaki, et al. (2003). "Analysis of cell lineage in two- and four-cell mouse embryos." Development **130**(21): 5113-22.
Fukuda, T., K. Kugou, et al. (2008). "Targeted induction of meiotic double-strand breaks reveals chromosomal domain-dependent regulation of Spo11 and interactions among potential sites of meiotic recombination." Nucleic Acids Res **36**(3): 984-97.
function, m. f. a. (1997). mRNA formation and function. New York, Academic Press.
Futterer, A., A. Raya, et al. (2011). "Ablation of Dido3 compromises lineage commitment of stem cells in vitro and during early embryonic development." Cell Death Differ.
Garcia-Cruz, R., A. Casanovas, et al. (2010). "Cytogenetic analyses of human oocytes provide new data on non-disjunction mechanisms and the origin of trisomy 16." Hum Reprod **25**(1): 179-91.
Gardner, R. L. (1997). "The early blastocyst is bilaterally symmetrical and its axis of symmetry is aligned with the animal-vegetal axis of the zygote in the mouse." Development **124**(2): 289-301.
Gardner, R. L. (2001). "Specification of embryonic axes begins before cleavage in normal mouse development." Development **128**(6): 839-47.
Gatti, J. L., S. Castella, et al. (2004). "Post-testicular sperm environment and fertility." Anim Reprod Sci **82-83**: 321-39.
Gaudet, F., W. M. Rideout, 3rd, et al. (2004). "Dnmt1 expression in pre- and postimplantation embryogenesis and the maintenance of IAP silencing." Mol Cell Biol **24**(4): 1640-8.
Gavrieli, Y., Y. Sherman, et al. (1992). "Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation." J Cell Biol **119**(3): 493-501.
Gearhart, J. D., M. T. Davisson, et al. (1986). "Autosomal aneuploidy in mice: generation and developmental consequences." Brain Res Bull **16**(6): 789-801.
Gearhart, J. D., M. L. Oster-Granite, et al. (1987). "Developmental consequences of autosomal aneuploidy in mammals." Dev Genet **8**(4): 249-65.
Geiman, T. M., L. Tessarollo, et al. (2001). "Lsh, a SNF2 family member, is required for normal murine development." Biochim Biophys Acta **1526**(2): 211-20.
Ginsburg, M., M. H. Snow, et al. (1990). "Primordial germ cells in the mouse embryo during gastrulation." Development **110**(2): 521-8.
Giordano, J., Y. Ge, et al. (2007). "Evolutionary history of mammalian transposons determined by genome-wide defragmentation." PLoS Comput Biol **3**(7): e137.
Giordano, J., Y. Ge, et al. (2007). "Evolutionary history of mammalian transposons determined by genome-wide defragmentation." PLoS Comput Biol **3**(7): e137.
Giorgini, F., H. G. Davies, et al. (2002). "Translational repression by MSY4 inhibits spermatid differentiation in mice." Development **129**(15): 3669-79.
Girard, A., R. Sachidanandam, et al. (2006). "A germline-specific class of small RNAs binds mammalian Piwi proteins." Nature **442**(7099): 199-202.
Gray, N. K., J. M. Coller, et al. (2000). "Multiple portions of poly(A)-binding protein stimulate translation in vivo." Embo J **19**(17): 4723-33.
Greenbaum, M. P., N. Iwamori, et al. (2009). "Mouse TEX14 is required for embryonic germ cell intercellular bridges but not female fertility." Biol Reprod **80**(3): 449-57.
Greenbaum, M. P., W. Yan, et al. (2006). "TEX14 is essential for intercellular bridges and fertility in male mice." Proc Natl Acad Sci U S A **103**(13): 4982-7.
Gropp, A. (1975). "Proceedings: Hypoplasia and malformation in trisomy syndromes of the mouse foetus." Mutat Res **29**(2): 216.
Gropp, A., H. Winking, et al. (1983). "Murine trisomy: developmental profiles of the embryo, and isolation of trisomic cellular systems." J Exp Zool **228**(2): 253-69.
Gubbay, J., J. Collignon, et al. (1990). "A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes." Nature **346**(6281): 245-50.
Hajkova, P., S. Erhardt, et al. (2002). "Epigenetic reprogramming in mouse primordial germ cells." Mech Dev **117**(1-2): 15-23.
Hale, D. W., L. L. Washburn, et al. (1993). "Meiotic abnormalities in hybrid mice of the C57BL/6J x Mus spretus cross suggest a cytogenetic basis for Haldane's rule of hybrid sterility." Cytogenet Cell Genet **63**(4): 221-34.

Hall, H. E., E. R. Chan, et al. (2007a). "The origin of trisomy 13." Am J Med Genet A **143A**(19): 2242-8.
Hall, H. E., U. Surti, et al. (2007b). "The origin of trisomy 22: evidence for acrocentric chromosome-specific patterns of nondisjunction." Am J Med Genet A **143A**(19): 2249-55.
Hamatani, T., M. G. Carter, et al. (2004a). "Dynamics of global gene expression changes during mouse preimplantation development." Dev Cell **6**(1): 117-31.
Hamatani, T., T. Daikoku, et al. (2004b). "Global gene expression analysis identifies molecular pathways distinguishing blastocyst dormancy and activation." Proc Natl Acad Sci U S A **101**(28): 10326-31.
Hamer, G., K. Gell, et al. (2006). "Characterization of a novel meiosis-specific protein within the central element of the synaptonemal complex." J Cell Sci **119**(Pt 19): 4025-32.
Hamer, G., I. Novak, et al. (2008). "Disruption of pairing and synapsis of chromosomes causes stage-specific apoptosis of male meiotic cells." Theriogenology **69**(3): 333-9.
Hammarstedt, M., K. Wallengren, et al. (2000). "Minimal exclusion of plasma membrane proteins during retroviral envelope formation." Proc Natl Acad Sci U S A **97**(13): 7527-32.
Handel, M. A. and J. C. Schimenti (2010). "Genetics of mammalian meiosis: regulation, dynamics and impact on fertility." Nat Rev Genet **11**(2): 124-36.
Hara, M. R., N. Agrawal, et al. (2005). "S-nitrosylated GAPDH initiates apoptotic cell death by nuclear translocation following Siah1 binding." Nat Cell Biol **7**(7): 665-74.
Hart, A. H., L. Hartley, et al. (2004). "Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human." Dev Dyn **230**(1): 187-98.
Hassold, T. and D. Chiu (1985). "Maternal age-specific rates of numerical chromosome abnormalities with special reference to trisomy." Hum Genet **70**(1): 11-7.
Hassold, T., H. Hall, et al. (2007). "The origin of human aneuploidy: where we have been, where we are going." Hum Mol Genet **16 Spec No. 2**: R203-8.
Hassold, T. and P. Hunt (2001). "To err (meiotically) is human: the genesis of human aneuploidy." Nat Rev Genet **2**(4): 280-91.
Hata, K., M. Kusumi, et al. (2006). "Meiotic and epigenetic aberrations in Dnmt3L-deficient male germ cells." Mol Reprod Dev **73**(1): 116-22.
Hauf, S., E. Roitinger, et al. (2005). "Dissociation of cohesin from chromosome arms and loss of arm cohesion during early mitosis depends on phosphorylation of SA2." PLoS Biol **3**(3): e69.
Hauf, S. and Y. Watanabe (2004). "Kinetochore orientation in mitosis and meiosis." Cell **119**(3): 317-27.
Hawkins, S. M. and M. M. Matzuk (2008). "The menstrual cycle: basic biology." Ann N Y Acad Sci **1135**: 10-8.
Hayashi, K., S. M. de Sousa Lopes, et al. (2007). "Germ cell specification in mice." Science **316**(5823): 394-6.
Hernandez, D. and E. M. Fisher (1999). "Mouse autosomal trisomy: two's company, three's a crowd." Trends Genet **15**(6): 241-7.
Hillman, N., M. I. Sherman, et al. (1972). "The effect of spatial arrangement on cell determination during mouse development." J Embryol Exp Morphol **28**(2): 263-78.
Hilscher, B., W. Hilscher, et al. (1974). "Kinetics of gametogenesis. I. Comparative histological and autoradiographic studies of oocytes and transitional prospermatogonia during oogenesis and prespermatogenesis." Cell Tissue Res **154**(4): 443-70.
Hirano, T. (2005). "SMC proteins and chromosome mechanics: from bacteria to humans." Philos Trans R Soc Lond B Biol Sci **360**(1455): 507-14.
Hirshfield, A. N. (1991). "Development of follicles in the mammalian ovary." Int Rev Cytol **124**: 43-101.
Hodges, C. A., E. Revenkova, et al. (2005). "SMC1beta-deficient female mice provide evidence that cohesins are a missing link in age-related nondisjunction." Nat Genet **37**(12): 1351-5.
Hogan, B. L., R. Beddington, et al. (1994). Manipulating the Mouse Embryo Cold Spring Harbor Laboratory Press.
Homer, H. A., A. McDougall, et al. (2005). "Mad2 prevents aneuploidy and premature proteolysis of cyclin B and securin during meiosis I in mouse oocytes." Genes Dev **19**(2): 202-7.
Howlett, S. K. and V. N. Bolton (1985). "Sequence and regulation of morphological and molecular events during the first cell cycle of mouse embryogenesis." J Embryol Exp Morphol **87**: 175-206.

http://scientopia.org/blogs/scicurious/2010/03/10/basics-guest-post-2-spermatogenesis/. (2011). "Spermatogenesis." Retrieved 22.09., 2011.
Huckins, C. and Y. Clermont (1968). "Evolution of gonocytes in the rat testis during late embryonic and early post-natal life." Arch Anat Histol Embryol **51**(1): 341-54.
Huda, A., L. Marino-Ramirez, et al. (2010). "Epigenetic histone modifications of human transposable elements: genome defense versus exaptation." Mob DNA **1**(1): 2.
Hunt, P. A. and T. J. Hassold (2002). "Sex matters in meiosis." Science **296**(5576): 2181-3.
Hunter, N. and N. Kleckner (2001). "The single-end invasion: an asymmetric intermediate at the double-strand break to double-holliday junction transition of meiotic recombination." Cell **106**(1): 59-70.
Hutnick, L. K., X. Huang, et al. (2010). "Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation-independent mechanism." J Biol Chem **285**(27): 21082-91.
Hyafil, F., D. Morello, et al. (1980). "A cell surface glycoprotein involved in the compaction of embryonal carcinoma cells and cleavage stage embryos." Cell **21**(3): 927-34.
Jackson, R. J., C. U. Hellen, et al. (2010). "The mechanism of eukaryotic translation initiation and principles of its regulation." Nat Rev Mol Cell Biol **11**(2): 113-27.
Jackson-Grusby, L., C. Beard, et al. (2001). "Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation." Nat Genet **27**(1): 31-9.
Jagarlamudi, K., P. Reddy, et al. (2010). "Genetically modified mouse models for premature ovarian failure (POF)." Mol Cell Endocrinol **315**(1-2): 1-10.
Jallepalli, P. V. and C. Lengauer (2001). "Chromosome segregation and cancer: cutting through the mystery." Nat Rev Cancer **1**(2): 109-17.
Jin, F., M. Hamada, et al. (2010). "Cdc20 is critical for meiosis I and fertility of female mice." PLoS Genet **6**(9).
Johnson, M. H. (2009). "From mouse egg to mouse embryo: polarities, axes, and tissues." Annu Rev Cell Dev Biol **25**: 483-512.
Johnson, M. H. and J. M. McConnell (2004). "Lineage allocation and cell polarity during mouse embryogenesis." Semin Cell Dev Biol **15**(5): 583-97.
Johnson, M. H. and C. A. Ziomek (1981a). "The foundation of two distinct cell lineages within the mouse morula." Cell **24**(1): 71-80.
Johnson, M. H. and C. A. Ziomek (1981b). "Induction of polarity in mouse 8-cell blastomeres: specificity, geometry, and stability." J Cell Biol **91**(1): 303-8.
Johnson, M. H. and C. A. Ziomek (1983). "Cell interactions influence the fate of mouse blastomeres undergoing the transition from the 16- to the 32-cell stage." Dev Biol **95**(1): 211-8.
Jones, K. T. (2008). "Meiosis in oocytes: predisposition to aneuploidy and its increased incidence with age." Hum Reprod Update **14**(2): 143-58.
Jones, P. A. and G. Liang (2009). "Rethinking how DNA methylation patterns are maintained." Nat Rev Genet **10**(11): 805-11.
Jurka, J., V. V. Kapitonov, et al. (2007). "Repetitive sequences in complex genomes: structure and evolution." Annu Rev Genomics Hum Genet **8**: 241-59.
Kalousek, D. K. (1991). "Pathology of abortion: chromosomal and genetic correlations." Monogr Pathol(33): 228-56.
Kanellopoulou, C., S. A. Muljo, et al. (2005). "Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing." Genes Dev **19**(4): 489-501.
Kano, H., I. Godoy, et al. (2009). "L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism." Genes Dev **23**(11): 1303-12.
Karimi, M. M., P. Goyal, et al. (2011). "DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs." Cell Stem Cell **8**(6): 676-87.
Kato, A., Y. Nagata, et al. (2004). "Delta-tubulin is a component of intercellular bridges and both the early and mature perinuclear rings during spermatogenesis." Dev Biol **269**(1): 196-205.
Kaufman, M. H., E. Huberman, et al. (1975). "Genetic control of haploid parthenogenetic development in mammalian embryos." Nature **254**(5502): 694-5.

Kelly, S. J. (1977). "Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres." J Exp Zool **200**(3): 365-76.
Kim, Y., A. Kobayashi, et al. (2006). "Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination." PLoS Biol **4**(6): e187.
Kimura-Yoshida, C., H. Nakano, et al. (2005). "Canonical Wnt signaling and its antagonist regulate anterior-posterior axis polarization by guiding cell migration in mouse visceral endoderm." Dev Cell **9**(5): 639-50.
Kishigami, S. and T. Wakayama (2007). "Efficient strontium-induced activation of mouse oocytes in standard culture media by chelating calcium." J Reprod Dev **53**(6): 1207-15.
Klein, F., P. Mahr, et al. (1999). "A central role for cohesins in sister chromatid cohesion, formation of axial elements, and recombination during yeast meiosis." Cell **98**(1): 91-103.
Kneitz, B., P. E. Cohen, et al. (2000). "MutS homolog 4 localization to meiotic chromosomes is required for chromosome pairing during meiosis in male and female mice." Genes Dev **14**(9): 1085-97.
Kocer, A., J. Reichmann, et al. (2009). "Germ cell sex determination in mammals." Mol Hum Reprod **15**(4): 205-13.
Koehler, K. E., R. S. Hawley, et al. (1996). "Recombination and nondisjunction in humans and flies." Hum Mol Genet **5 Spec No**: 1495-504.
Koehler, K. E., S. E. Schrump, et al. (2006). "Near-human aneuploidy levels in female mice with homeologous chromosomes." Curr Biol **16**(15): R579-80.
Koopman, P., J. Gubbay, et al. (1991). "Male development of chromosomally female mice transgenic for Sry." Nature **351**(6322): 117-21.
Kouznetsova, A., L. Lister, et al. (2007). "Bi-orientation of achiasmatic chromosomes in meiosis I oocytes contributes to aneuploidy in mice." Nat Genet **39**(8): 966-8.
Krishnan, S., S. Horowitz, et al. (2011). "Structure and function of histone H3 lysine 9 methyltransferases and demethylases." Chembiochem **12**(2): 254-63.
Kudo, N. R., M. Anger, et al. (2009). "Role of cleavage by separase of the Rec8 kleisin subunit of cohesin during mammalian meiosis I." J Cell Sci **122**(Pt 15): 2686-98.
Kudo, N. R., K. Wassmann, et al. (2006). "Resolution of chiasmata in oocytes requires separase-mediated proteolysis." Cell **126**(1): 135-46.
Kumaki, Y., M. Oda, et al. (2008). "QUMA: quantification tool for methylation analysis." Nucleic Acids Res **36**(Web Server issue): W170-5.
Kuramochi-Miyagawa, S., T. Kimura, et al. (2004). "Mili, a mammalian member of piwi family gene, is essential for spermatogenesis." Development **131**(4): 839-49.
Kuramochi-Miyagawa, S., T. Kimura, et al. (2001). "Two mouse piwi-related genes: miwi and mili." Mech Dev **108**(1-2): 121-33.
Kuramochi-Miyagawa, S., T. Watanabe, et al. (2008). "DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes." Genes Dev **22**(7): 908-17.
Kurotaki, Y., K. Hatta, et al. (2007). "Blastocyst axis is specified independently of early cell lineage but aligns with the ZP shape." Science **316**(5825): 719-23.
Kwon, G. S., M. Viotti, et al. (2008). "The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages." Dev Cell **15**(4): 509-20.
Kwon, Y. T., Z. Xia, et al. (2003). "Female lethality and apoptosis of spermatocytes in mice lacking the UBR2 ubiquitin ligase of the N-end rule pathway." Mol Cell Biol **23**(22): 8255-71.
Kwon, Y. T., Z. Xia, et al. (2001). "Construction and analysis of mouse strains lacking the ubiquitin ligase UBR1 (E3alpha) of the N-end rule pathway." Mol Cell Biol **21**(23): 8007-21.
Lamb, N. E., S. L. Sherman, et al. (2005). "Effect of meiotic recombination on the production of aneuploid gametes in humans." Cytogenet Genome Res **111**(3-4): 250-5.
Lande-Diner, L., J. Zhang, et al. (2007). "Role of DNA methylation in stable gene repression." J Biol Chem **282**(16): 12194-200.
Latham, K. E. and R. M. Schultz (2001). "Embryonic genome activation." Front Biosci **6**: D748-59.
Latham, K. E., D. Solter, et al. (1992). "Acquisition of a transcriptionally permissive state during the 1-cell stage of mouse embryogenesis." Dev Biol **149**(2): 457-62.

Lawson, K. A., N. R. Dunn, et al. (1999). "Bmp4 is required for the generation of primordial germ cells in the mouse embryo." Genes Dev **13**(4): 424-36.
Lawson, K. A., J. J. Meneses, et al. (1991). "Clonal analysis of epiblast fate during germ layer formation in the mouse embryo." Development **113**(3): 891-911.
Lee, J., T. S. Kitajima, et al. (2008). "Unified mode of centromeric protection by shugoshin in mammalian oocytes and somatic cells." Nat Cell Biol **10**(1): 42-52.
Lees-Murdock, D. J., M. De Felici, et al. (2003). "Methylation dynamics of repetitive DNA elements in the mouse germ cell lineage." Genomics **82**(2): 230-7.
LeMaire-Adkins, R., K. Radke, et al. (1997). "Lack of checkpoint control at the metaphase/anaphase transition: a mechanism of meiotic nondisjunction in mammalian females." J Cell Biol **139**(7): 1611-9.
Li, E. (2002). "Chromatin modification and epigenetic reprogramming in mammalian development." Nat Rev Genet **3**(9): 662-73.
Li, E., C. Beard, et al. (1993). "Role for DNA methylation in genomic imprinting." Nature **366**(6453): 362-5.
Li, E., T. H. Bestor, et al. (1992). "Targeted mutation of the DNA methyltransferase gene results in embryonic lethality." Cell **69**(6): 915-26.
Li, P. W., J. Li, et al. (2006). "The dicistronic RNA from the mouse LINE-1 retrotransposon contains an internal ribosome entry site upstream of each ORF: implications for retrotransposition." Nucleic Acids Res **34**(3): 853-64.
Li, W. and H. Ma (2006). "Double-stranded DNA breaks and gene functions in recombination and meiosis." Cell Res **16**(5): 402-12.
Li, X. C. and J. C. Schimenti (2007). "Mouse pachytene checkpoint 2 (trip13) is required for completing meiotic recombination but not synapsis." PLoS Genet **3**(8): e130.
Libby, B. J., R. De La Fuente, et al. (2002). "The mouse meiotic mutation mei1 disrupts chromosome synapsis with sexually dimorphic consequences for meiotic progression." Dev Biol **242**(2): 174-87.
Lightfoot, D. A., A. Kouznetsova, et al. (2006). "The fate of mosaic aneuploid embryos during mouse development." Dev Biol **289**(2): 384-94.
Lister, L. M., A. Kouznetsova, et al. (2010). "Age-related meiotic segregation errors in mammalian oocytes are preceded by depletion of cohesin and Sgo2." Curr Biol **20**(17): 1511-21.
Liu, P., M. Wakamiya, et al. (1999). "Requirement for Wnt3 in vertebrate axis formation." Nat Genet **22**(4): 361-5.
Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-8.
Llano, E., R. Gomez, et al. (2008). "Shugoshin-2 is essential for the completion of meiosis but not for mitotic cell division in mice." Genes Dev **22**(17): 2400-13.
Longo, L., A. Bygrave, et al. (1997). "The chromosome make-up of mouse embryonic stem cells is predictive of somatic and germ cell chimaerism." Transgenic Res **6**(5): 321-8.
Lorke, D. E. (1994). "Developmental characteristics of trisomy 19 mice." Acta Anat (Basel) **150**(3): 159-69.
Losada, A., M. Hirano, et al. (1998). "Identification of Xenopus SMC protein complexes required for sister chromatid cohesion." Genes Dev **12**(13): 1986-97.
Lotz, K., G. Pyrowolakis, et al. (2004). "BRUCE, a giant E2/E3 ubiquitin ligase and inhibitor of apoptosis protein of the trans-Golgi network, is required for normal placenta development and mouse survival." Mol Cell Biol **24**(21): 9339-50.
Lovell-Badge, R. and E. Robertson (1990). "XY female mice resulting from a heritable mutation in the primary testis-determining gene, Tdy." Development **109**(3): 635-46.
Lucifero, D., S. La Salle, et al. (2007). "Coordinate regulation of DNA methyltransferase expression during oogenesis." BMC Dev Biol **7**: 36.
Lucifero, D., M. R. Mann, et al. (2004). "Gene-specific timing and epigenetic memory in oocyte imprinting." Hum Mol Genet **13**(8): 839-49.
Luoh, S. W., P. A. Bain, et al. (1997). "Zfx mutation results in small animal size and reduced germ cell number in male and female mice." Development **124**(11): 2275-84.

Ma, L., G. M. Buchold, et al. (2009). "GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline." PLoS Genet **5**(9): e1000635.
Maatouk, D. M., L. DiNapoli, et al. (2008). "Stabilization of beta-catenin in XY gonads causes male-to-female sex-reversal." Hum Mol Genet **17**(19): 2949-55.
Maksakova, I. A., D. L. Mager, et al. (2008). "Keeping active endogenous retroviral-like elements in check: the epigenetic perspective." Cell Mol Life Sci **65**(21): 3329-47.
Maksakova, I. A., M. T. Romanish, et al. (2006). "Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line." PLoS Genet **2**(1): e2.
Malik, H. S., W. D. Burke, et al. (1999). "The age and evolution of non-LTR retrotransposable elements." Mol Biol Evol **16**(6): 793-805.
Maratou, K., T. Forster, et al. (2004). "Expression profiling of the developing testis in wild-type and Dazl knockout mice." Mol Reprod Dev **67**(1): 26-54.
Marston, A. L. and A. Amon (2004). "Meiosis: cell-cycle controls shuffle and deal." Nat Rev Mol Cell Biol **5**(12): 983-97.
Martin, S. L. (1991). "LINEs." Curr Opin Genet Dev **1**(4): 505-8.
Matson, C. K., M. W. Murphy, et al. (2011). "DMRT1 prevents female reprogramming in the postnatal mammalian testis." Nature **476**(7358): 101-4.
Matzuk, M. M. and D. J. Lamb (2002). "Genetic dissection of mammalian fertility pathways." Nat Cell Biol **4 Suppl**: s41-9.
Matzuk, M. M. and D. J. Lamb (2008). "The biology of infertility: research advances and clinical challenges." Nat Med **14**(11): 1197-213.
McClintock, B. (1956). "Controlling elements and the gene." Cold Spring Harb Symp Quant Biol **21**: 197-216.
McGee, E. A. and A. J. Hsueh (2000). "Initial and cyclic recruitment of ovarian follicles." Endocr Rev **21**(2): 200-14.
McGuinness, B. E., T. Hirota, et al. (2005). "Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells." PLoS Biol **3**(3): e86.
McLaren, A. (1984). "Meiosis and differentiation of mouse germ cells." Symp Soc Exp Biol **38**: 7-23.
McLaren, A. (2003). "Primordial germ cells in the mouse." Dev Biol **262**(1): 1-15.
McLaren, A. and D. Southee (1997). "Entry of mouse embryonic germ cells into meiosis." Dev Biol **187**(1): 107-13.
Mets, D. G. and B. J. Meyer (2009). "Condensins regulate meiotic DNA break distribution, thus crossover frequency, by controlling chromosome structure." Cell **139**(1): 73-86.
Mitsui, K., Y. Tokuzawa, et al. (2003). "The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells." Cell **113**(5): 631-42.
Moens, P. B., E. Marcon, et al. (2007). "Initiation and resolution of interhomolog connections: crossover and non-crossover sites along mouse synaptonemal complexes." J Cell Sci **120**(Pt 6): 1017-27.
Mogk, A., R. Schmidt, et al. (2007). "The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies." Trends Cell Biol **17**(4): 165-72.
Moody, S. A. (1999). Cell Lineage and Fate Determination, ACADEMIC PRESS.
Morelli, M. A. and P. E. Cohen (2005). "Not all germ cells are created equal: aspects of sexual dimorphism in mammalian meiosis." Reproduction **130**(6): 761-81.
Morriss-Kay, G. and B. Putz (1986). "Abnormal neural fold development in mouse trisomy 12 and trisomy 14. II. LM and TEM." Brain Res Bull **16**(6): 825-32.
Motosugi, N., T. Bauer, et al. (2005). "Polarity of the mouse embryo is established at blastocyst and is not prepatterned." Genes Dev **19**(9): 1081-92.
Mulnard, J. G. (1992). "The Brussels School of Embryology." Int J Dev Biol **36**(1): 17-24.
Munne, S., C. Magli, et al. (1999). "Positive outcome after preimplantation diagnosis of aneuploidy in human embryos." Hum Reprod **14**(9): 2191-9.
Munne, S., M. Sandalinas, et al. (2002). "Chromosome mosaicism in cleavage-stage human embryos: evidence of a maternal age effect." Reprod Biomed Online **4**(3): 223-32.
Muotri, A. R., M. C. Marchetto, et al. (2010). "L1 retrotransposition in neurons is modulated by MeCP2." Nature **468**(7322): 443-6.

Murchison, E. P., P. Stein, et al. (2007). "Critical roles for Dicer in the female germline." Genes Dev **21**(6): 682-93.
Nakagawa, T., Y. Nabeshima, et al. (2007). "Functional identification of the actual and potential stem cell compartments in mouse spermatogenesis." Dev Cell **12**(2): 195-206.
Nasmyth, K. and C. H. Haering (2005). "The structure and function of SMC and kleisin complexes." Annu Rev Biochem **74**: 595-648.
Ng, R. K., W. Dean, et al. (2008). "Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5." Nat Cell Biol **10**(11): 1280-90.
Nichols, J., E. P. Evans, et al. (1990). "Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity." Development **110**(4): 1341-8.
Nichols, J. and A. Smith (2009). "Naive and primed pluripotent states." Cell Stem Cell **4**(6): 487-92.
Nichols, J., B. Zevnik, et al. (1998). "Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4." Cell **95**(3): 379-91.
Ninomiya, T., M. Vuillemin, et al. (1993). "Mouse fetal trisomy 13 and hypotrophy of the spinal cord: effect on calbindin-D28k and calretinin expressed by neurons of the spinal cord and dorsal root ganglia." Neuroscience **57**(4): 1109-20.
Nishioka, N., S. Yamamoto, et al. (2008). "Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos." Mech Dev **125**(3-4): 270-83.
Niwa, H., J. Miyazaki, et al. (2000). "Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells." Nat Genet **24**(4): 372-6.
Niwa, H., Y. Toyooka, et al. (2005). "Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation." Cell **123**(5): 917-29.
Nothias, J. Y., S. Majumder, et al. (1995). "Regulation of gene expression at the beginning of mammalian development." J Biol Chem **270**(38): 22077-80.
Oatley, J. M. and R. L. Brinster (2008). "Regulation of spermatogonial stem cell self-renewal in mammals." Annu Rev Cell Dev Biol **24**: 263-86.
O'Donnell, L., P. K. Nicholls, et al. (2011). "Spermiation: The process of sperm release." Spermatogenesis **1**(1): 14-35.
Ohinata, Y., B. Payer, et al. (2005). "Blimp1 is a critical determinant of the germ cell lineage in mice." Nature **436**(7048): 207-13.
Ollinger, R., M. Alsheimer, et al. (2005). "Mammalian protein SCP1 forms synaptonemal complex-like structures in the absence of meiotic chromosomes." Mol Biol Cell **16**(1): 212-7.
Ollinger, R., A. J. Childs, et al. (2008). "Deletion of the pluripotency-associated Tex19.1 gene causes activation of endogenous retroviruses and defective spermatogenesis in mice." PLoS Genet **4**(9): e1000199.
Ollinger, R., J. Reichmann, et al. (2010). "Meiosis and retrotransposon silencing during germ cell development in mice." Differentiation **79**(3): 147-58.
Orr-Weaver, T. (1996). "Meiotic nondisjunction does the two-step." Nat Genet **14**(4): 374-6.
Orwig, K. E., B. Y. Ryu, et al. (2002). "Male germ-line stem cell potential is predicted by morphology of cells in neonatal rat testes." Proc Natl Acad Sci U S A **99**(18): 11706-11.
Ouyang, Y., Y. T. Kwon, et al. (2006). "Loss of Ubr2, an E3 ubiquitin ligase, leads to chromosome fragility and impaired homologous recombinational repair." Mutat Res **596**(1-2): 64-75.
Pacchierotti, F., I. D. Adler, et al. (2007). "Gender effects on the incidence of aneuploidy in mammalian germ cells." Environ Res **104**(1): 46-69.
Page, S. L. and R. S. Hawley (2003). "Chromosome choreography: the meiotic ballet." Science **301**(5634): 785-9.
Page, S. L. and R. S. Hawley (2004). "The genetics and molecular biology of the synaptonemal complex." Annu Rev Cell Dev Biol **20**: 525-58.
Palmer, S. J. and P. S. Burgoyne (1991). "In situ analysis of fetal, prepuberal and adult XX----XY chimaeric mouse testes: Sertoli cells are predominantly, but not exclusively, XY." Development **112**(1): 265-8.
Palmieri, S. L., W. Peter, et al. (1994). "Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation." Dev Biol **166**(1): 259-67.

Pandur, P. D., S. A. Sullivan, et al. (2002). "Multiple maternal influences on dorsal-ventral fate of Xenopus animal blastomeres." Dev Dyn **225**(4): 581-7.
Paques, F. and J. E. Haber (1999). "Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae." Microbiol Mol Biol Rev **63**(2): 349-404.
Patel, K., J. Dickson, et al. (2010). "Targeting of 5-aza-2'-deoxycytidine residues by chromatin-associated DNMT1 induces proteasomal degradation of the free enzyme." Nucleic Acids Res **38**(13): 4313-24.
Peaston, A. E., A. V. Evsikov, et al. (2004). "Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos." Dev Cell **7**(4): 597-606.
Pellestor, F., T. Anahory, et al. (2005). "Effect of maternal age on the frequency of cytogenetic abnormalities in human oocytes." Cytogenet Genome Res **111**(3-4): 206-12.
Pellestor, F., B. Andreo, et al. (2002). "Mechanisms of non-disjunction in human female meiosis: the co-existence of two modes of malsegregation evidenced by the karyotyping of 1397 in-vitro unfertilized oocytes." Hum Reprod **17**(8): 2134-45.
Pellestor, F., B. Andreo, et al. (2003). "Maternal aging and chromosomal abnormalities: new data drawn from in vitro unfertilized human oocytes." Hum Genet **112**(2): 195-203.
Pelttari, J., M. R. Hoja, et al. (2001). "A meiotic chromosomal core consisting of cohesin complex proteins recruits DNA recombination proteins and promotes synapsis in the absence of an axial element in mammalian meiotic cells." Mol Cell Biol **21**(16): 5667-77.
Pelttari, J., M. R. Hoja, et al. (2001). "A meiotic chromosomal core consisting of cohesin complex proteins recruits DNA recombination proteins and promotes synapsis in the absence of an axial element in mammalian meiotic cells." Mol Cell Biol **21**(16): 5667-77.
Pepling, M. E. and A. C. Spradling (1998). "Female mouse germ cells form synchronously dividing cysts." Development **125**(17): 3323-8.
Pepling, M. E. and A. C. Spradling (2001). "Mouse ovarian germ cell cysts undergo programmed breakdown to form primordial follicles." Dev Biol **234**(2): 339-51.
Perea-Gomez, A., F. D. Vella, et al. (2002). "Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks." Dev Cell **3**(5): 745-56.
Petronczki, M., M. F. Siomos, et al. (2003). "Un menage a quatre: the molecular biology of chromosome segregation in meiosis." Cell **112**(4): 423-40.
Pezzi, N., I. Prieto, et al. (2000). "STAG3, a novel gene encoding a protein involved in meiotic chromosome pairing and location of STAG3-related genes flanking the Williams-Beuren syndrome deletion." Faseb J **14**(3): 581-92.
Piotrowska, K., F. Wianny, et al. (2001). "Blastomeres arising from the first cleavage division have distinguishable fates in normal mouse development." Development **128**(19): 3739-48.
Piotrowska-Nitsche, K., A. Perea-Gomez, et al. (2005a). "Four-cell stage mouse blastomeres have different developmental properties." Development **132**(3): 479-90.
Piotrowska-Nitsche, K. and M. Zernicka-Goetz (2005b). "Spatial arrangement of individual 4-cell stage blastomeres and the order in which they are generated correlate with blastocyst pattern in the mouse embryo." Mech Dev **122**(4): 487-500.
Pittman, D. L., J. Cobb, et al. (1998). "Meiotic prophase arrest with failure of chromosome synapsis in mice deficient for Dmc1, a germline-specific RecA homolog." Mol Cell **1**(5): 697-705.
Plusa, B., A. K. Hadjantonakis, et al. (2005). "The first cleavage of the mouse zygote predicts the blastocyst axis." Nature **434**(7031): 391-5.
Prieto, I., N. Pezzi, et al. (2002). "STAG2 and Rad21 mammalian mitotic cohesins are implicated in meiosis." EMBO Rep **3**(6): 543-50.
Prieto, I., J. A. Suja, et al. (2001). "Mammalian STAG3 is a cohesin specific to sister chromatid arms in meiosis I." Nat Cell Biol **3**(8): 761-6.
Puchtler, H., S. N. Meloan, et al. (1986). "Application of current chemical concepts to metal-hematein and -brazilein stains." Histochemistry **85**(5): 353-64.
Ralston, A. and J. Rossant (2008). "Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo." Dev Biol **313**(2): 614-29.
Rao, H., F. Uhlmann, et al. (2001). "Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability." Nature **410**(6831): 955-9.

Rashbass, P., L. A. Cooke, et al. (1991). "A cell autonomous function of Brachyury in T/T embryonic stem cell chimaeras." Nature **353**(6342): 348-51.
Rebuzzini, P., T. Neri, et al. (2008). "Karyotype analysis of the euploid cell population of a mouse embryonic stem cell line revealed a high incidence of chromosome abnormalities that varied during culture." Cytogenet Genome Res **121**(1): 18-24.
Renfree, M. B. and G. Shaw (2000). "Diapause." Annu Rev Physiol **62**: 353-75.
Requena, A., F. Bronet, et al. (2009). "The impact of in-vitro maturation of oocytes on aneuploidy rate." Reprod Biomed Online **18**(6): 777-83.
Reuss, F. U. (1992). "Expression of intracisternal A-particle-related retroviral element-encoded envelope proteins detected in cell lines." J Virol **66**(4): 1915-23.
Reuter, M., S. Chuma, et al. (2009). "Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile." Nat Struct Mol Biol **16**(6): 639-46.
Revenkova, E., M. Eijpe, et al. (2001). "Novel meiosis-specific isoform of mammalian SMC1." Mol Cell Biol **21**(20): 6984-98.
Revenkova, E., M. Eijpe, et al. (2004). "Cohesin SMC1 beta is required for meiotic chromosome dynamics, sister chromatid cohesion and DNA recombination." Nat Cell Biol **6**(6): 555-62.
Revenkova, E. and R. Jessberger (2005). "Keeping sister chromatids together: cohesins in meiosis." Reproduction **130**(6): 783-90.
Reynolds, N., B. Collier, et al. (2005). "Dazl binds in vivo to specific transcripts and can regulate the pre-meiotic translation of Mvh in germ cells." Hum Mol Genet **14**(24): 3899-909.
Richardson, L., M. E. Torres-Padilla, et al. (2006). "Regionalised signalling within the extraembryonic ectoderm regulates anterior visceral endoderm positioning in the mouse embryo." Mech Dev **123**(4): 288-96.
Rivera-Perez, J. A., J. Mager, et al. (2003). "Dynamic morphogenetic events characterize the mouse visceral endoderm." Dev Biol **261**(2): 470-87.
Robinson, W. P., B. D. Kuchinka, et al. (1998). "Maternal meiosis I non-disjunction of chromosome 15: dependence of the maternal age effect on level of recombination." Hum Mol Genet **7**(6): 1011-9.
Rodriguez, T. A., S. Srinivas, et al. (2005). "Induction and migration of the anterior visceral endoderm is regulated by the extra-embryonic ectoderm." Development **132**(11): 2513-20.
Romanienko, P. J. and R. D. Camerini-Otero (2000). "The mouse Spo11 gene is required for meiotic chromosome synapsis." Mol Cell **6**(5): 975-87.
Romanish, M. T., C. J. Cohen, et al. (2010
). "Potential mechanisms of endogenous retroviral-mediated genomic instability in human cancer." Semin Cancer Biol **20**(4): 246-53.
Roper, R. J. and R. H. Reeves (2006). "Understanding the basis for Down syndrome phenotypes." PLoS Genet **2**(3): e50.
Rosenbusch, B. (2004). "The incidence of aneuploidy in human oocytes assessed by conventional cytogenetic analysis." Hereditas **141**(2): 97-105.
Rosner, M. H., M. A. Vigano, et al. (1990). "A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo." Nature **345**(6277): 686-92.
Rossant, J. (1976). "Postimplantation development of blastomeres isolated from 4- and 8-cell mouse eggs." J Embryol Exp Morphol **36**(2): 283-90.
Rossant, J. and P. P. Tam (2009). "Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse." Development **136**(5): 701-13.
Ruggiu, M., R. Speed, et al. (1997). "The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis." Nature **389**(6646): 73-7.
Russ, A. P., S. Wattler, et al. (2000). "Eomesodermin is required for mouse trophoblast development and mesoderm formation." Nature **404**(6773): 95-9.
Russell, L. D., H. P. Ren, et al. (1990). "A comparative study in twelve mammalian species of volume densities, volumes, and numerical densities of selected testis components, emphasizing those related to the Sertoli cell." Am J Anat **188**(1): 21-30.
Sachs, A. B. and R. W. Davis (1989). "The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation." Cell **58**(5): 857-67.

Saitou, M. (2009). "Specification of the germ cell lineage in mice." Front Biosci **14**: 1068-87.
Saitou, M., S. C. Barton, et al. (2002). "A molecular programme for the specification of germ cell fate in mice." Nature **418**(6895): 293-300.
Sakuno, T., K. Tada, et al. (2009). "Kinetochore geometry defined by cohesion within the centromere." Nature **458**(7240): 852-8.
Sambrook, R. (2001). Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press.
Sandalinas, M., S. Sadowy, et al. (2001). "Developmental ability of chromosomally abnormal human embryos to develop to the blastocyst stage." Hum Reprod **16**(9): 1954-8.
Sasaki, H. (2010). "Mechanisms of trophectoderm fate specification in preimplantation mouse development." Dev Growth Differ **52**(3): 263-73.
Sasaki, H. and Y. Matsui (2008). "Epigenetic events in mammalian germ-cell development: reprogramming and beyond." Nat Rev Genet **9**(2): 129-40.
Schalk, J. A., A. J. Dietrich, et al. (1998). "Localization of SCP2 and SCP3 protein molecules within synaptonemal complexes of the rat." Chromosoma **107**(8): 540-8.
Schier, A. F. and W. S. Talbot (2005). "Molecular genetics of axis formation in zebrafish." Annu Rev Genet **39**: 561-613.
Scholer, H. R., G. R. Dressler, et al. (1990). "Oct-4: a germline-specific transcription factor mapping to the mouse t-complex." Embo J **9**(7): 2185-95.
Schrans-Stassen, B. H., H. J. van de Kant, et al. (1999). "Differential expression of c-kit in mouse undifferentiated and differentiating type A spermatogonia." Endocrinology **140**(12): 5894-900.
Schrans-Stassen, B. H., H. J. van de Kant, et al. (1999). "Differential expression of c-kit in mouse undifferentiated and differentiating type A spermatogonia." Endocrinology **140**(12): 5894-900.
Schultz, R. M. (2002). "The molecular foundations of the maternal to zygotic transition in the preimplantation embryo." Hum Reprod Update **8**(4): 323-31.
Schvarzstein, M., S. M. Wignall, et al. (2010). "Coordinating cohesion, co-orientation, and congression during meiosis: lessons from holocentric chromosomes." Genes Dev **24**(3): 219-28.
Schwacha, A. and N. Kleckner (1994). "Identification of joint molecules that form frequently between homologs but rarely between sister chromatids during yeast meiosis." Cell **76**(1): 51-63.
Schwacha, A. and N. Kleckner (1997). "Interhomolog bias during meiotic recombination: meiotic functions promote a highly differentiated interhomolog-only pathway." Cell **90**(6): 1123-35.
Sears, D. D., J. H. Hegemann, et al. (1992). "Meiotic recombination and segregation of human-derived artificial chromosomes in Saccharomyces cerevisiae." Proc Natl Acad Sci U S A **89**(12): 5296-300.
Seki, Y., K. Hayashi, et al. (2005). "Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice." Dev Biol **278**(2): 440-58.
Seki, Y., M. Yamaji, et al. (2007). "Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice." Development **134**(14): 2627-38.
Sekido, R., I. Bar, et al. (2004). "SOX9 is up-regulated by the transient expression of SRY specifically in Sertoli cell precursors." Dev Biol **274**(2): 271-9.
Senner, C. E. and M. Hemberger (2010). "Regulation of early trophoblast differentiation - lessons from the mouse." Placenta **31**(11): 944-50.
Seydoux, G. and R. E. Braun (2006). "Pathway to totipotency: lessons from germ cells." Cell **127**(5): 891-904.
Sherman, S. L., N. E. Lamb, et al. (2006). "Relationship of recombination patterns and maternal age among non-disjoined chromosomes 21." Biochem Soc Trans **34**(Pt 4): 578-80.
Shinohara, T. and R. L. Brinster (2000). "Enrichment and transplantation of spermatogonial stem cells." Int J Androl **23 Suppl 2**: 89-91.
Shintomi, K. and T. Hirano (2009). "Releasing cohesin from chromosome arms in early mitosis: opposing actions of Wapl-Pds5 and Sgo1." Genes Dev **23**(18): 2224-36.
Shintomi, K. and T. Hirano (2010). "Sister chromatid resolution: a cohesin releasing network and beyond." Chromosoma **119**(5): 459-67.

Shirayoshi, Y., T. S. Okada, et al. (1983). "The calcium-dependent cell-cell adhesion system regulates inner cell mass formation and cell surface polarization in early mouse development." Cell **35**(3 Pt 2): 631-8.
Shoji, M., T. Tanaka, et al. (2009). "The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline." Dev Cell **17**(6): 775-87.
Shyu, A. B., M. F. Wilkinson, et al. (2008). "Messenger RNA regulation: to translate or to degrade." Embo J **27**(3): 471-81.
Silva, J., O. Barrandon, et al. (2008). "Promotion of reprogramming to ground state pluripotency by signal inhibition." PLoS Biol **6**(10): e253.
Sinclair, A. H., P. Berta, et al. (1990). "A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif." Nature **346**(6281): 240-4.
Smagulova, F., I. V. Gregoretti, et al. (2010). "Genome-wide analysis reveals novel molecular features of mouse recombination hotspots." Nature **472**(7343): 375-8.
Smith, B. S. and N. J. Walker (1992). "Lens and mesenchyme ultrastructure in gestational day 11 trisomy 1 mice." Teratology **46**(6): 583-98.
Soper, S. F., G. W. van der Heijden, et al. (2008). "Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis." Dev Cell **15**(2): 285-97.
Soudais, C., M. Bielinska, et al. (1995). "Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation in vitro." Development **121**(11): 3877-88.
Souter, I., L. M. Baltagi, et al. (2010). "Women, weight, and fertility: the effect of body mass index on the outcome of superovulation/intrauterine insemination cycles." Fertil Steril **95**(3): 1042-7.
Speed, R. M. (1982). "Meiosis in the foetal mouse ovary. I. An analysis at the light microscope level using surface-spreading." Chromosoma **85**(3): 427-37.
Speek, M. (2001). "Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes." Mol Cell Biol **21**(6): 1973-85.
Spencer, R. J., B. C. Del Rosario, et al. (2011). "A Boundary Element Between Tsix and Xist Binds the Chromatin Insulator Ctcf and Contributes to Initiation of X Chromosome Inactivation." Genetics.
Sripathy, S. P., J. Stevens, et al. (2006). "The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression." Mol Cell Biol **26**(22): 8623-38.
Steinberger, E., K. D. Smith, et al. (1979). "Testosterone levels in female partners of infertile couples. Relationship between androgen levels in the woman, the male factor, and the incidence of pregnancy." Am J Obstet Gynecol **133**(2): 133-8.
Stitzel, M. L. and G. Seydoux (2007). "Regulation of the oocyte-to-zygote transition." Science **316**(5823): 407-8.
Stoye, J. P. (2001). "Endogenous retroviruses: still active after all these years?" Curr Biol **11**(22): R914-6.
Strumpf, D., C. A. Mao, et al. (2005). "Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst." Development **132**(9): 2093-102.
Sugawara, A., K. Goto, et al. (2006). "Current status of chromosomal abnormalities in mouse embryonic stem cell lines used in Japan." Comp Med **56**(1): 31-4.
Sugimoto, M. and K. Abe (2007). "X chromosome reactivation initiates in nascent primordial germ cells in mice." PLoS Genet **3**(7): e116.
Sullivan, M. and D. O. Morgan (2007). "Finishing mitosis, one step at a time." Nat Rev Mol Cell Biol **8**(11): 894-903.
Sun, L. Q., D. W. Lee, et al. (2004). "Growth retardation and premature aging phenotypes in mice with disruption of the SNF2-like gene, PASG." Genes Dev **18**(9): 1035-46.
Surani, M. A. (2001). "Reprogramming of genome function through epigenetic inheritance." Nature **414**(6859): 122-8.
Suwinska, A., R. Czolowska, et al. (2008). "Blastomeres of the mouse embryo lose totipotency after the fifth cleavage division: expression of Cdx2 and Oct4 and developmental potential of inner and outer blastomeres of 16- and 32-cell embryos." Dev Biol **322**(1): 133-44.
Symington, L. S. and J. Gautier (2011). "Double-Strand Break End Resection and Repair Pathway Choice." Annu Rev Genet.

Tajima, K., M. Orisaka, et al. (2007). "Ovarian theca cells in follicular function." Reprod Biomed Online **15**(5): 591-609.
Tam, P. P. and D. A. Loebel (2007). "Gene function in mouse embryogenesis: get set for gastrulation." Nat Rev Genet **8**(5): 368-81.
Tam, P. P. and M. H. Snow (1981). "Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos." J Embryol Exp Morphol **64**: 133-47.
Tam, P. P. and S. X. Zhou (1996). "The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo." Dev Biol **178**(1): 124-32.
Tarkowski, A. K. (1959). "Experiments on the development of isolated blastomers of mouse eggs." Nature **184**: 1286-7.
Tarkowski, A. K. and J. Wroblewska (1967). "Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage." J Embryol Exp Morphol **18**(1): 155-80.
Tasaki, T. and Y. T. Kwon (2007). "The mammalian N-end rule pathway: new insights into its components and physiological roles." Trends Biochem Sci **32**(11): 520-8.
Tasaki, T., L. C. Mulder, et al. (2005). "A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons." Mol Cell Biol **25**(16): 7120-36.
Tasaki, T., R. Sohr, et al. (2007b). "Biochemical and genetic studies of UBR3, a ubiquitin ligase with a function in olfactory and other sensory systems." J Biol Chem **282**(25): 18510-20.
Tasaki, T., A. Zakrzewska, et al. (2009). "The substrate recognition domains of the N-end rule pathway." J Biol Chem **284**(3): 1884-95.
Thomas, N. S., S. Ennis, et al. (2001). "Maternal sex chromosome non-disjunction: evidence for X chromosome-specific risk factors." Hum Mol Genet **10**(3): 243-50.
Thomas, P. and R. Beddington (1996). "Anterior primitive endoderm may be responsible for patterning the anterior neural plate in the mouse embryo." Curr Biol **6**(11): 1487-96.
Torres-Padilla, M. E., L. Richardson, et al. (2007). "The anterior visceral endoderm of the mouse embryo is established from both preimplantation precursor cells and by de novo gene expression after implantation." Dev Biol **309**(1): 97-112.
Toth, A., K. P. Rabitsch, et al. (2000). "Functional genomics identifies monopolin: a kinetochore protein required for segregation of homologs during meiosis i." Cell **103**(7): 1155-68.
Uhlenhaut, N. H., S. Jakob, et al. (2009). "Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation." Cell **139**(6): 1130-42.
Uhlmann, F., F. Lottspeich, et al. (1999). "Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1." Nature **400**(6739): 37-42.
Uhlmann, F. and K. Nasmyth (1998). "Cohesion between sister chromatids must be established during DNA replication." Curr Biol **8**(20): 1095-101.
Uhlmann, F., D. Wernic, et al. (2000). "Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast." Cell **103**(3): 375-86.
Unhavaithaya, Y., Y. Hao, et al. (2009). "MILI, a PIWI-interacting RNA-binding protein, is required for germ line stem cell self-renewal and appears to positively regulate translation." J Biol Chem **284**(10): 6507-19.
URL, R.-H. W. D. and http://www.repeatmasker.org/ [Accessed June 12 (Smit, A.F.A., Hubley, R., Green, P.,).
van Echten-Arends, J., S. Mastenbroek, et al. (2011). "Chromosomal mosaicism in human preimplantation embryos: a systematic review." Hum Reprod Update **17**(5): 620-7.
Vanderhyden, B. (2002). "Molecular basis of ovarian development and function." Front Biosci **7**: d2006-22.
Vanneste, E., T. Voet, et al. (2009). "Chromosome instability is common in human cleavage-stage embryos." Nat Med **15**(5): 577-83.
Varshavsky, A. (1997). "The N-end rule pathway of protein degradation." Genes Cells **2**(1): 13-28.
Varshavsky, A. (2011). "The N-end rule pathway and regulation by proteolysis." Protein Sci.
Verdaasdonk, J. S. and K. Bloom (2011). "Centromeres: unique chromatin structures that drive chromosome segregation." Nat Rev Mol Cell Biol **12**(5): 320-32.

Vestweber, D., A. Gossler, et al. (1987). "Expression and distribution of cell adhesion molecule uvomorulin in mouse preimplantation embryos." Dev Biol **124**(2): 451-6.
Vialard, F., C. Petit, et al. (2006). "Evidence of a high proportion of premature unbalanced separation of sister chromatids in the first polar bodies of women of advanced age." Hum Reprod **21**(5): 1172-8.
Vidal, V. P., M. C. Chaboissier, et al. (2001). "Sox9 induces testis development in XX transgenic mice." Nat Genet **28**(3): 216-7.
Waizenegger, I. C., S. Hauf, et al. (2000). "Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase." Cell **103**(3): 399-410.
Wallace, N., B. J. Wagstaff, et al. (2008). "LINE-1 ORF1 protein enhances Alu SINE retrotransposition." Gene **419**(1-2): 1-6.
Walsh, C. P. and T. H. Bestor (1999). "Cytosine methylation and mammalian development." Genes Dev **13**(1): 26-34.
Walsh, C. P., J. R. Chaillet, et al. (1998). "Transcription of IAP endogenous retroviruses is constrained by cytosine methylation." Nat Genet **20**(2): 116-7.
Wang, H. and S. K. Dey (2006). "Roadmap to embryo implantation: clues from mouse models." Nat Rev Genet **7**(3): 185-99.
Watanabe, T., Y. Totoki, et al. (2008). "Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes." Nature **453**(7194): 539-43.
Watson, E. D. and J. C. Cross (2005). "Development of structures and transport functions in the mouse placenta." Physiology (Bethesda) **20**: 180-93.
Wells, D. and J. D. Delhanty (2000). "Comprehensive chromosomal analysis of human preimplantation embryos using whole genome amplification and single cell comparative genomic hybridization." Mol Hum Reprod **6**(11): 1055-62.
Wells, D., J. K. Sherlock, et al. (1999). "Detailed chromosomal and molecular genetic analysis of single cells by whole genome amplification and comparative genomic hybridisation." Nucleic Acids Res **27**(4): 1214-8.
White, J. A. and J. Heasman (2008). "Maternal control of pattern formation in Xenopus laevis." J Exp Zool B Mol Dev Evol **310**(1): 73-84.
Wolstenholme, J. and R. R. Angell (2000). "Maternal age and trisomy--a unifying mechanism of formation." Chromosoma **109**(7): 435-8.
Wozniak, R. J., W. T. Klimecki, et al. (2007). "5-Aza-2'-deoxycytidine-mediated reductions in G9A histone methyltransferase and histone H3 K9 di-methylation levels are linked to tumor suppressor gene reactivation." Oncogene **26**(1): 77-90.
Xu, H., M. D. Beasley, et al. (2005). "Absence of mouse REC8 cohesin promotes synapsis of sister chromatids in meiosis." Dev Cell **8**(6): 949-61.
Xu, M., Y. You, et al. (2008). "Mice deficient for a small cluster of Piwi-interacting RNAs implicate Piwi-interacting RNAs in transposon control." Biol Reprod **79**(1): 51-7.
Yagi, R., M. J. Kohn, et al. (2007). "Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development." Development **134**(21): 3827-36.
Yamaguchi, S., H. Kimura, et al. (2005). "Nanog expression in mouse germ cell development." Gene Expr Patterns **5**(5): 639-46.
Yamamoto, Y. and M. Oelgeschlager (2004). "Regulation of bone morphogenetic proteins in early embryonic development." Naturwissenschaften **91**(11): 519-34.
Yang, F., Y. Cheng, et al. (2010). "The ubiquitin ligase Ubr2, a recognition E3 component of the N-end rule pathway, stabilizes Tex19.1 during spermatogenesis." PLoS One **5**(11): e14017.
Yang, F., R. De La Fuente, et al. (2006). "Mouse SYCP2 is required for synaptonemal complex assembly and chromosomal synapsis during male meiosis." J Cell Biol **173**(4): 497-507.
Yates, A. and I. Chambers (2005). "The homeodomain protein Nanog and pluripotency in mouse embryonic stem cells." Biochem Soc Trans **33**(Pt 6): 1518-21.
Yeom, Y. I., G. Fuhrmann, et al. (1996). "Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells." Development **122**(3): 881-94.
Ying, Y., X. M. Liu, et al. (2000). "Requirement of Bmp8b for the generation of primordial germ cells in the mouse." Mol Endocrinol **14**(7): 1053-63.

Ying, Y. and G. Q. Zhao (2001). "Cooperation of endoderm-derived BMP2 and extraembryonic ectoderm-derived BMP4 in primordial germ cell generation in the mouse." Dev Biol **232**(2): 484-92.
Yoder, J. A., C. P. Walsh, et al. (1997). "Cytosine methylation and the ecology of intragenomic parasites." Trends Genet **13**(8): 335-40.
Yoder, J. A., C. P. Walsh, et al. (1997). "Cytosine methylation and the ecology of intragenomic parasites." Trends Genet **13**(8): 335-40.
Yokobayashi, S., M. Yamamoto, et al. (2003). "Cohesins determine the attachment manner of kinetochores to spindle microtubules at meiosis I in fission yeast." Mol Cell Biol **23**(11): 3965-73.
Yoshida, M., M. Kijima, et al. (1990). "Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A." J Biol Chem **265**(28): 17174-9.
Yoshida, S., M. Sukeno, et al. (2006). "The first round of mouse spermatogenesis is a distinctive program that lacks the self-renewing spermatogonia stage." Development **133**(8): 1495-505.
Yoshinaga, K., S. Nishikawa, et al. (1991). "Role of c-kit in mouse spermatogenesis: identification of spermatogonia as a specific site of c-kit expression and function." Development **113**(2): 689-99.
Youds, J. L. and S. J. Boulton (2011). "The choice in meiosis - defining the factors that influence crossover or non-crossover formation." J Cell Sci **124**(Pt 4): 501-13.
Yuan, L., J. G. Liu, et al. (2002). "Female germ cell aneuploidy and embryo death in mice lacking the meiosis-specific protein SCP3." Science **296**(5570): 1115-8.
Yuan, L., J. G. Liu, et al. (2000). "The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility." Mol Cell **5**(1): 73-83.
Zenker, M., J. Mayerle, et al. (2005). "Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome)." Nat Genet **37**(12): 1345-50.
Zernicka-Goetz, M. (1998). "Fertile offspring derived from mammalian eggs lacking either animal or vegetal poles." Development **125**(23): 4803-8.
Zernicka-Goetz, M. (2005). "Cleavage pattern and emerging asymmetry of the mouse embryo." Nat Rev Mol Cell Biol **6**(12): 919-28.
Zheng, K., J. Xiol, et al. (2010). "Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway." Proc Natl Acad Sci U S A **107**(26): 11841-6.
Zvetkova, I., A. Apedaile, et al. (2005). "Global hypomethylation of the genome in XX embryonic stem cells." Nat Genet **37**(11): 1274-9.

APPENDIX

Microarray Analysis of LTR Retrotransposon Silencing Identifies Hdac1 as a Regulator of Retrotransposon Expression in Mouse Embryonic Stem Cells

Judith Reichmann[1,#], James Crichton[1,#], Monika J Madej[1], Mary Taggart[1], Philippe Gautier[1], Jose Luis Garcia-Perez[2], Richard R Meehan[1] and Ian R Adams[1*]

[1]MRC Human Genetics Unit at the Institute of Genetics and Molecular Medicine at the University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU. United Kingdom.

[2]GENYO, Pfizer-University of Granada-Andalusian Government-Centre for Genomics and Oncological Research, Avda Ilustración 114, Granada 18007, Spain

[#] These authors contributed equally to this work

[*] Corresponding author (Ian.Adams@igmm.ed.ac.uk)

Abstract

Retrotransposons are highly prevalent in mammalian genomes due to their ability to amplify in pluripotent cells or developing germ cells. Host mechanisms that silence retrotransposons in germ cells and pluripotent cells are important for limiting the accumulation of the repetitive elements in the genome during evolution. However, although silencing of selected individual retrotransposons can be relatively well-studied, many mammalian retrotransposons are seldom analysed and their silencing in germ cells, pluripotent cells or somatic cells remains poorly understood. Here we show, and experimentally verify, that cryptic repetitive element probes present in Illumina and Affymetrix gene expression microarray platforms can accurately and sensitively monitor repetitive element expression data. This computational approach to genome-wide retrotransposon expression has allowed us to identify the histone deacetylase Hdac1 as a component of the retrotransposon silencing machinery in mouse embryonic stem cells, and to determine the retrotransposon targets of Hdac1 in these cells. We also identify retrotransposons that are targets of other retrotransposon silencing mechanisms such as DNA methylation, Eset-mediated histone modification, and Ring1B/Eed-containing polycomb repressive complexes in mouse embryonic stem cells. Furthermore, our computational analysis of retrotransposon silencing suggests that multiple silencing mechanisms are independently targeted to retrotransposons in embryonic stem cells, that different genomic copies of the same retrotransposon can be differentially sensitive to these silencing mechanisms, and helps define retrotransposon sequence elements that are targeted by silencing machineries. Thus repeat annotation of gene expression microarray data suggests that a complex interplay between silencing mechanisms represses retrotransposon loci in germ cells and embryonic stem cells.

Author Summary

Repetitive DNA sequences make up almost half the mammalian genome. A large proportion of mammalian repetitive DNA sequences use RNA intermediates to amplify and insert themselves into new locations in the genome. Mammalian genomes contain hundreds of different types of these mutagenic retrotransposons, but the mechanisms that host cells use to silence most of these elements are poorly understood. Here we describe a computational approach to monitor expression of hundreds of different retrotransposons in gene expression microarray datasets. This approach reveals new retrotransposon targets for silencing mechanisms such as DNA methylation, histone modification and polycomb repression in mouse embryonic stem cells, and identifies the histone deacetylase Hdac1 as a regulator of retrotransposons in this cell type. These computational predictions are verified experimentally by qRT-PCR in *Dnmt1*$^{-/-}$ *Dnmt3a*$^{-/-}$ *Dnmt3b*$^{-/-}$ embryonic stem cells, *Ring1B*$^{-/-}$ embryonic stem cells and *Hdac1*$^{-/-}$ embryonic stem cells. We also use microarray analysis of retrotransposon expression to show that the pluripotency-associated *Tex19.1* gene has exquisite specificity for *MMERVK10C* elements in developing male germ cells. Importantly, our computational analysis also suggests that different genomic copies of individual retrotransposons can be differentially regulated, and helps identify the sequences in these retrotransposons that are being targeted by the host cell's silencing mechanisms.

Introduction

Repetitive DNA sequences account for around forty per cent of sequenced mammalian genomes [1,2]. The most basic repetitive elements in mammalian genomes are tandem arrays of repeated monomeric DNA sequences. These simple repeats and satellite sequences have repeating units of around 1-5 bp and 100-500 bp respectively [3]. More complex classes of repetitive element include DNA transposons and retrotransposons, mobile genetic elements that are able to integrate into new sites in the genome. DNA transposons typically encode a transposase enzyme that catalyses the non-replicative mobilization of the DNA transposon through a cut and paste mechanism [4]. In contrast, retrotransposons mobilize using a replicative copy and paste mechanism that involves an RNA intermediate. However this retrotransposition can occur by fundamentally different mechanisms depending on the structure of the retrotransposon [5,6]. DNA transposons and retrotransposons account for ~0.9% and ~37% of the mouse genome respectively [2]. However, while DNA transposon activity appears to be extinct in the mouse genome, retrotransposons remain active [2]. Mouse retrotransposons include long interspersed elements (*LINE*s), short interspersed elements (*SINE*s), and long terminal repeat (LTR) retrotransposons [3]. Full-length class I *LINE*s are ~7 kb long and encode two proteins that are required for the reverse-transcription of *LINE-1* RNA and its subsequent integration into new sites in the genome [7]. *SINE*s are derived from reverse-transcription of small cellular RNAs and utilise *LINE-1* proteins *in trans* to mediate retrotransposition [8]. LTR retrotransposons, also known as endogenous retroviruses (ERVs), either encode *gag*, *pol*, *pro* and sometimes also *env* genes, or use the retroviral genes encoded by other ERVs, to drive a retroviral life-cycle [2,3,9].

Retrotransposons have the potential to alter the genomic landscape and change gene expression when they amplify or integrate into new sites in the host genome, providing an important driving force for evolutionary change [10]. Although retrotransposition can occur in somatic cells [11,12], repetitive elements need to amplify in germ cells, or their pluripotent precursors, in order to

successfully propagate. The Repeatmasker database of repetitive elements [13] currently contains consensus sequences for 1221 different types of repetitive element, each of which is present in multiple copies in the mouse genome. These 1221 repetitive elements are organized into 16 different classes comprising a total of 45 families (see Supporting Figure S1 for a schematic overview of this organization). The repetitive element classes that contain the greatest number of different repetitive elements are LTR retrotransposons (471 elements), simple repeats (315 elements), DNA transposons (156 elements) and *LINE* retrotransposons (122 elements). Many of the repetitive elements that are present in the mammalian genome are poorly characterized, and it is often not clear whether different elements within each class or family are active at similar stages of germ cell or pluripotent cell development, or whether different elements are recognized and regulated by the same host defence mechanisms. Indeed the rich diversity of successful repetitive elements in the mammalian genome may indicate that different elements have evolved different strategies to evade recognition or suppression by host defence mechanisms.

The high mutational load associated with excessive amplification of repetitive elements in the developing germline is likely to be detrimental to the evolutionary success of the host organism. Much progress has been made in identifying and understanding the mechanisms that suppress the activity of repetitive elements in germ cells and pluripotent cells, particularly transcriptional repression of retrotransposon activity in mice [reviewed in 14–17]. Epigenetic modifications such as DNA methylation, histone methylation and histone deacetylation are all implicated in transcriptional silencing of retrotransposons. DNA methylation is required for transcriptional repression of intracisternal A particle (*IAP*) elements, a member of the ERVK family of LTR retrotransposons, in somatic cells and germ cells [18,19]. Targeting DNA methylation to *IAP* elements during male fetal germ cell development requires the interaction between the piwi-piRNA pathway and DNA methyltransferase enzymes [reviewed in 15–17]. In pluripotent cells such as embryonic stem (ES) cells, mutations in all three catalytically active DNA methyltransferases

greatly reduce the levels of DNA methylation in the genome [20], and these *Dnmt1*$^{-/-}$ *Dnmt3a*$^{-/-}$ *Dnmt3b*$^{-/-}$ triple knock out (Dnmt TKO) ES cells have increased expression of *IAP* retrotransposons [21,22]. However, the increase in *IAP* expression in Dnmt TKO ES cells is relatively modest compared to somatic cells, and ES cells appear to rely more on the transcriptional co-repressor Kap1 to repress *IAP* elements [21–23]. Kap1 probably acts through recruitment of histone H3K9 methyltransferases, primarily Eset (also known as Setdb1 or Kmt1e), to deposit repressive histone modifications on *IAP* chromatin [22,23]. Together Kap1 and Eset have been shown to target various ERV1, ERVK and ERVL LTR retrotransposons [22–24]. However, different silencing mechanisms are likely to be operating on retrotransposons that are not enriched for H3K9 methylation in mouse ES cells [16,25]. Polycomb repressive complex (PRC)-mediated H3K27 trimethylation and Lsd1-dependent H3K4 demethylation are also implicated in transcriptional repression of LTR retrotransposons in mouse ES cells [26,27], and histone deacetylation has been implicated in transcriptional silencing of newly-integrated *LINE-1* elements in undifferentiated human embryonal carcinoma (EC) cells [28]. Histone deacetylases, DNA methyltransferases, histone lysine methyltransferases and PRC proteins are all also implicated in transcriptional silencing of retroviral LTRs in human somatic cells [e.g. 29,30], and some of the mechanisms operating to repress retrotransposon transcription in somatic cells may operate in pluripotent cells too. In addition to transcriptional silencing, retrotransposon activity is also regulated at post-transcriptional levels in germ cells and pluripotent cells through the activity of miRNAs and endogenous small interfering RNAs (endo-siRNAs) [31–33]. Other host factors, such as Apobec proteins [34] and the Trex1 endonuclease [35], have been shown to suppress retrotransposon activity post-transcriptionally in somatic cell types, and similar factors presumably also operate in pluripotent cells [36] and germ cells. Thus, multiple mechanisms probably combine to bring about effective silencing of different classes of retrotransposon in different cell types.

Although silencing of repetitive elements has been studied by qRT-PCR and Northern blotting of representative candidate elements in ES cells and in other cell types, few genome-wide studies of the repetitive element expression have been performed to date [22,23,37]. Therefore it is often not clear how many different repetitive elements are being targeted by a specific silencing mechanism in any particular cell type. Given the antagonistic evolutionary relationship between retrotransposon expression and host silencing mechanisms, identifying repetitive elements that have escaped specific host silencing mechanisms may generate some insight into how these mechanisms are able to determine which regions of the genome or transcriptome to target. Microarrays are widely used for gene expression profiling, and a large volume of microarray gene expression data obtained under various experimental conditions has been deposited in freely-accessible repositories such as NCBI GEO [38]. Microarray analysis of gene expression has been able to identify some changes in repetitive element gene expression [e.g. 26,39], but although a number of probes present on commercially available microarrays are identical to repetitive element sequences, few probes on these arrays are explicitly annotated as recognising repetitive elements.

The purpose of this study is to computationally extract information about genome-wide silencing of repetitive elements in germ cells and stem cells from microarray gene expression data. Using this approach we identify retrotransposons that are silenced by DNA methylation and various histone modifications in mouse embryonic stem cells. We also identify the histone deacetylase Hdac1 as a regulator of retrotransposons in mouse ES cells. Our results demonstrate that different silencing mechanisms can be independently recruited to retrotransposons in a modular manner, and that different genomic copies of individual retrotransposons can be differentially sensitive to loss of these silencing mechanisms. Lastly we show that analysing the sequence variation between differentially regulated copies of individual retrotransposons can help identify sequences important for retrotransposon silencing.

Results

Identification of Repetitive Element Probes in the Illumina and Affymetrix Gene Expression Microarray Platforms

Previously, in a study designed to refine and improve the detection of gene expression changes in Illumina Mouse WG-6 Beadchip microarrays data, more than 4,000 probes in the Illumina Mouse WG-6 Beadchips were identified that map to regions of the mouse genome that are at least partially masked by Repeatmasker [40]. Although information from these probes was discarded from gene expression microarray data in that study in order to improve the analysis of the remaining single-copy probes [40], these repeat probes could potentially contain information about genome-wide repetitive element expression in microarray datasets. We therefore investigated how well different classes of repetitive element are represented in Illumina Beadarrays, and whether these probes could monitor repetitive element expression on a genome-wide level.

The Illumina Mouse WG-6 Beadchips each contain ~46,000 probes. We identified ~2,300 repetitive element probes in version 1.0, version 1.1 and version 2.0 of these arrays (Table 1) by comparing the genomic locations of the probes with the Repeatmasked regions of the mouse genome (see Materials and Methods). The proportion of repetitive element probes identified on the Illumina Beadchips in this analysis (~5%) is around half that reported previously [40]. This difference appears to be a consequence of using stricter criteria to identify repetitive element probes in the current study. In each version of the Illumina Mouse WG-6 Beadchip analyzed, ~1400 probes were in the correct orientation to detect sense repetitive element transcripts. Text files containing the repetitive element probe names and sequences identified in the Illumina Mouse WG-6 Beadchip are included online (Supporting Datasets S2-S4). Of the 1221 different repetitive elements in the mouse genome annotated in the Repeatmasker database, ~320 are represented by probes in the different versions of the Illumina Mouse WG-6 Beadchips (Table 2). Repetitive elements belonging to the *LINE* and *SINE* classes are well represented on these arrays, and repetitive elements belonging to the LTR retrotransposon and DNA transposon classes are reasonably represented (Table 2). Simple

repeats and satellite repeats are also present but less well represented on the Illumina Mouse WG-6 Beadchips (Table 2). Thus Illumina Mouse WG-6 Beadchips have a good coverage of probes for monitoring transposon and retrotransposon expression during genome-wide transcriptional profiling.

We applied the same rationale to identify repetitive element probes present in the Affymetrix Murine Genome U74Av2 and Mouse Expression 430 2.0 GeneChips (Table 1). The Murine Genome U74Av2 and Mouse Expression 430 2.0 GeneChips contain ~4,200 and ~26,000 probes respectively that are in the correct orientation to detect sense transcripts from repetitive elements. Text files containing the repetitive element probe names and sequences identified in the Affymetrix Gene Expression GeneChips are included online (Supporting Dataset S5, S6). Like the Illumina Mouse WG-6 Beadchip arrays, the Affymetrix arrays also have good representation of repetitive elements belonging to *LINE* and *SINE* classes, and the Affymetrix Mouse Expression 430 2.0 GeneChip also has good coverage of LTR retrotransposons and DNA transposons (Table 2). The Affymetrix Murine Genome U74Av2 GeneChip has reasonable coverage of repetitive elements within the LTR retrotransposon and DNA transposon class (Table 2). Thus Affymetrix Gene Expression GeneChips also contain a wide range probes that can be used to monitor transposon and retrotransposon expression.

Computational Analysis of Repetitive Element Expression in *Tex19.1*$^{-/-}$ Testes From Microarray Gene Expression Profiles

We had previously identified upregulation of the *MMERVK10C* (ERVK family) LTR retrotransposon in mouse germ cells lacking the pluripotency-associated *Tex19.1*$^{-/-}$ gene by analysing individual probe sequences upregulated in Illumina Beadchip microarray data [39]. In order to test whether any additional retrotransposons might be targets for *Tex19.1* in developing male germ cells we used the repeat probes in the Illumina Mouse WG-6 v2.0 Beadchip to assess

genome-wide repetitive element expression in *Tex19.1*$^{-/-}$ testis microarray data. As *Tex19.1*$^{-/-}$ male mice have defects in progression through meiosis that perturb the normal cellular composition of the testis, gene expression profiling was performed on 16 dpp prepubertal testes undergoing the first wave of spermatogenesis where defects in meiosis are first becoming apparent [39]. In addition, the *Tex19.1* mutation was backcrossed onto an inbred C57BL/6 genetic background in order to minimize genetic variation between the animals used for this microarray analysis. 19,089 probes on the Illumina Beadarray were expressed in 16 dpp testes in this experiment (Figure 1A), with most showing no significant change in expression in *Tex19.1*$^{-/-}$ testes. The expression level of 158 probes (0.8%) are downregulated at least 2 fold in *Tex19.1*$^{-/-}$ testes at a significance level of $p<0.01$. However, the apparent downregulation of many of these probes may be a consequence of the delay in meiotic progression that is becoming evident in *Tex19.1*$^{-/-}$ testes at 16 dpp [39]. On the other hand, 10 probes (0.05%) are upregulated at least 2 fold in *Tex19.1*$^{-/-}$ testes at $p<0.01$.

In general the repetitive element probes behaved similarly to other probes on the array (Figure 1A). 512 (2.7%) of the 19,089 probes expressed in 16 dpp testes are repeat probes. These 512 repeat probes represent 173 different repetitive elements. LTR retrotransposon, *LINE*, *SINE*, DNA transposon, and satellite transcripts were all expressed similarly in *Tex19.1*$^{-/-}$ and control testes (Figure 1A). However, 6 repeat probes belonging to the LTR retrotransposon class appear to be behaving as outliers from the total probe population (Figure 1A). These outlying probes are upregulated 2-4 fold in *Tex19.1*$^{-/-}$ testes, and all belong to the ERVK family of LTR retrotransposons (Figure 1B). All of these 6 probes are complementary to the *MMERVK10C* repetitive element (Figure 1C). Indeed, although the 124 LTR retrotransposon probes that are expressed in this dataset do not behave differently from the 18,577 non-repeat probes (Figure 1D, Wilcoxon rank sum test $p=0.5$), the 9 *MMERVK10C* probes expressed in this dataset represent a distinct population from the non-repeat probes (Figure 1D, Wilcoxon rank sum test, $p<0.0001$). The *MMERVK10C* probes also appear to be behaving differently from other LTR retrotransposon and

ERVK retrotransposon probes in this dataset (Wilcoxon rank sum tests, $p<0.0001$). Only four non-repeat probes are upregulated in *Tex19.1*$^{-/-}$ testes, and none of these probes map close to *MMERVK10C* loci in the reference genome, suggesting that the upregulation of *MMERVK10C* elements in *Tex19.1*$^{-/-}$ testes is likely to be caused by loss of a *trans*-acting retrotransposon silencing mechanism rather than changes in non-repetitive gene expression affecting the local chromatin structure and influencing expression of nearby retrotransposon sequences.

The unique behaviour of *MMERVK10C* repeat probes in the microarray data was confirmed by identifying probes whose expression changed at least 2 fold ($p<0.01$) in *Tex19.1*$^{-/-}$ testes relative to control testes. 6 (1.2%) of the 512 repeat probes change expression at least 2 fold ($p<0.01$) in *Tex19.1*$^{-/-}$ testes, and all 6 of these repeat probes are derived from *MMERVK10C-int* LTR retrotransposon sequences. We confirmed that each of these *MMERVK10C* probe sequences matches multiple genomic loci (≥48/50 nt identity) by BLAT suggesting that each probe is able to detect expression from multiple genomic copies of the *MMERVK10C* LTR retrotransposon (data not shown). Furthermore, we also confirmed that the non-complementary repeat probes recognizing antisense repetitive element transcripts did not show any significant change in expression in *Tex19.1*$^{-/-}$ testes (data not shown). Thus repeat-annotation of the *Tex19.1*$^{-/-}$ Illumina Beadchip data suggests that expression of *MMERVK10C* retrotransposons is significantly and specifically upregulated in *Tex19.1*$^{-/-}$ testes. The systematic annotation and analysis of the C57BL/6 *Tex19.1*$^{-/-}$ testis microarray data presented here is consistent with our previous findings that *MMERVK10C* elements are upregulated in *Tex19.1*$^{-/-}$ testes from a mixed (129/Ola x CD1) genetic background [39], but importantly also extends the range and variety of repetitive elements analysed in these animals. Intriguingly, *MMERVK10C* remains the only repetitive element among the 173 elements represented in this dataset whose expression changes by more than 2 fold in the absence of *Tex19.1*.

Retrotransposon Derepression in *Tex19.1*$^{-/-}$ Testes is Restricted to *MMERVK10C* Elements

Our computational analysis of *Tex19.1*$^{-/-}$ testis microarray data suggests that repetitive element misexpression in *Tex19.1*$^{-/-}$ testes is largely restricted to upregulation of *MMERVK10C* elements (Figure 1A-D). We verified the upregulation of *MMERVK10C* elements in an independent group of C57BL/6 *Tex19.1*$^{-/-}$ testes by qRT-PCR (Figure 1E). The ~2 fold qRT-PCR upregulation of *MMERVK10C* elements in C57BL/6 *Tex19.1*$^{-/-}$ testes is similar to the ~4 fold qRT-PCR upregulation of this element reported previously using animals on a mixed genetic background [39]. The slightly lower level of upregulation of *MMERVK10C* seen in C57BL/6 animals may be caused by differences in the rate of testis development between these genetic backgrounds. In order to investigate the apparent specificity of the *MMERVK10C* upregulation evident in the microarray analysis we tested expression of *LINE-1* and some representative ERV1, ERVK and ERVL LTR retrotransposon sequences in *Tex19.1*$^{-/-}$ testes by qRT-PCR. qRT-PCR for *LINE-1* retrotransposons (Figure 1E) confirmed the repeat-annotation analysis suggesting that these elements do not change expression in *Tex19.1*$^{-/-}$ testes (Figure 1A-D). Furthermore, *RLTR4*, *ETnERV2* and *MERVL2a* LTR retrotransposons representing the ERV1, ERVK and ERVL families of LTR retrotransposons also do not change expression in *Tex19.1*$^{-/-}$ testes in either the Illumina Beadarray data (Figure 1A-D) or by qRT-PCR (Figure 1E). Thus *MMERVK10C* elements appear to be behaving differently from other LTR retrotransposons in *Tex19.1*$^{-/-}$ testes.

The Illumina Beadarrays used to profile gene expression in the *Tex19.1*$^{-/-}$ testes contain probes representing around a third of the LTR retrotransposons present in the mouse genome. Therefore although the computational and experimental data both suggest that *MMERVK10C* elements respond differently from other retrotransposons in the genome to the loss of *Tex19.1*, we investigated whether LTR retrotransposons that were closely related to *MMERVK10C* might also be upregulated in *Tex19.1*$^{-/-}$ testes. We used *MMERVK10C pol* and *pro* protein sequences to identify repetitive elements closely related to *MMERVK10C* (Figure 2A). *MMERVK10C* appears to be most closely related to *IAP* elements, with the *pol* protein sequences of *MMERVK10C*, *IAPEz* and

IAPEY3 all having around 75% similarity to each other. Although there are numerous *IAP* probes in the Illumina Beadarrays, these probes do not appear to be changing in *Tex19.1*$^{-/-}$ testes (Figure 2B). Furthermore we tested expression of *IAPEz* and *IAPEY3* elements in *Tex19.1*$^{-/-}$ testes by qRT-PCR (Figure 2C) and found that, as suggested by computational analysis of the microarray data, expression of these elements is not changing in *Tex19.1*$^{-/-}$ testes. We also tested expression of the *MMERVK9E* retrotransposon that is related to *MMERVK10C* but not represented on the Illumina Beadarrays. *MMERVK9E* has around 65% similarity to *MMERVK10C* across the *pol* protein sequence, but is not part of the cluster of *IAP* elements evident in the *MMERVK10C* phylogeny (Figure 2A). However, qRT-PCR data shows that *MMERVK9E* elements do not change expression in *Tex19.1*$^{-/-}$ testes either (Figure 2C). Thus retrotransposon derepression in *Tex19.1*$^{-/-}$ testes appears to be intriguingly restricted to *MMERVK10C* elements.

Different Transcriptional Silencing Mechanisms Have Distinct Effects on Genome-Wide Repression of Repetitive Elements

Our data on *Tex19.1*$^{-/-}$ testes suggests that only a small number of retrotransposon RNAs are sensitive to loss of *Tex19.1* in germ cells. We therefore next investigated whether loss of well established retrotransposon silencing mechanisms had more extensive effects on genome-wide repression of retrotransposons using ES cells as a model. We computationally analysed repetitive element expression in previously published gene expression microarray datasets from Dnmt TKO ES cells carrying mutations in all three catalytically active DNA methyltransferases [41], and from ES cells transiently transfected with shRNAs to knock-down the histone H3K9 methyltransferase Eset [42]. Although the Dnmt TKO and *Eset*^{shRNA} ES cell gene expression profiles were performed on Affymetrix and Illumina platforms respectively, and may therefore have some differences in coverage of individual retrotransposons or sensitivity of detection limits, different classes of repetitive elements are similarly represented on these platforms (Table 2) and some genome-wide comparisons will still be informative. We also included data from Affymetrix gene expression

profiling of ES cells carrying mutations in the Hdac1 histone deacetylase enzyme [43] in this analysis. Although the HDAC family of histone deacetylases are implicated in retrotransposon silencing by virtue of being targets of trichostatin A [28,44,45], the role and retrotransposon targets of the different HDAC histone deacetylases has not yet been defined. Genome-wide analysis of retrotransposon silencing in Dnmt TKO, *Eset*^{shRNA} and *Hdac1*$^{-/-}$ ES cells could therefore uncover new or additional retrotransposon targets for these mechanisms in ES cells.

Repeat-annotation of Dnmt TKO, *Eset*^{shRNA} and *Hdac1*$^{-/-}$ ES cells (Figure 3) confirmed that LTR retrotransposons are upregulated in all of these mutant ES cells. Interestingly, although individual retrotransposon sequences could be selected that show upregulation in each of these mutant ES cell lines, the genome-wide overview of retrotransposon behaviour shows striking differences in retrotransposon behaviour between mutant ES lines (Figure 3A, 3C, 3E). Dnmt TKO ES cells appear to modestly upregulate a number of LTR retrotransposon probes around 2-8 fold, which behave similarly to the upregulated non-repeat probes in the array, but other classes of repeat probe do not appear to change (Figure 3A). The upregulated group of LTR retrotransposon probes in Dnmt TKO ES cells is primarily composed of ERV1 and ERVK classes of LTR retrotransposon (Figure 3B). In contrast *Eset*^{shRNA} ES cells appear to strongly upregulate most LTR retrotransposon probes in the array, and these upregulated LTR retrotransposon probes appear to be responding more strongly to loss of *Eset* than the upregulated non-repeat probes in the dataset (Figure 3C). The range of LTR retrotransposon probes upregulated in *Eset*^{shRNA} ES cells is more expansive than in Dnmt TKO ES cells with probes belonging to ERV1, ERVK and ERVL classes all being upregulated (Figure 3D). Furthermore, *Eset*^{shRNA} ES cells appear to modestly upregulate *LINE-1* probes (Figure 3C), a group of retrotransposons that does not strongly change expression in Dnmt TKO ES cells (Figure 3A). Thus *Eset* appears to have a stronger and more widespread role in repressing retrotransposons in ES cells than DNA methylation. Interestingly, *Hdac1* also has a role in repressing retrotransposons in ES cells (Figure 3E). However the role of *Hdac1* appears to be

distinct from the roles of DNA methylation and Eset histone methyltransferase. *Hdac1*$^{-/-}$ ES cells upregulate one group of LTR retrotransposon probes 4-8 fold, a relatively strong upregulation compared to non-repeat probes in the dataset, and downregulate a second large group of LTR retrotransposon probes around 2-4 fold (Figure 3E). The upregulated and downregulated groups of LTR retrotransposon probes are both primarily composed of ERVK class LTR retrotransposons (Figure 3F, pink dots), and these changes in ERVK probe expression are comparable in magnitude to the changes in non-repetitive gene expression that occur in *Hdac1*$^{-/-}$ ES cells (Figure 3F, grey dots, [43]). The observation that LTR retrotransposon expression is altered in *Hdac1*$^{-/-}$ ES cells is consistent with data showing that human *HDAC1* can silence avian retroviral LTR reporter genes in somatic HeLa cells.[29,30], and identifies *Hdac1* as a novel regulator of retrotransposon expression in mouse ES cells. *Hdac1*$^{-/-}$ ES cells do not appear to change expression of other classes of repeat probe (Figure 3E), and therefore Hdac1 appears to be more restricted than either DNA methylation or Eset in the range of retrotransposon sequence classes that it affects. However unlike DNA methylation or Eset, Hdac can have both positive and negative effects on expression of retrotransposons. Thus although the Dnmt TKO, *Eset*^{shRNA}, and *Hdac1*$^{-/-}$ ES cell lines all upregulate individual retrotransposons, these mechanisms appear to have different effects on retrotransposon expression at a genome-wide level.

Interactions Between Retrotransposon Silencing Mechanisms in ES Cells

We next investigated how the Dnmt, Eset and Hdac1 transcriptional repression mechanisms interact in ES cells by identifying distinct and overlapping retrotransposon targets for these mechanisms. We identified repeat probes in each of the Dnmt TKO, *Eset*^{shRNA}, and *Hdac1*$^{-/-}$ ES cell datasets that changed expression at least 2 fold ($p<0.01$) relative to the appropriate wild-type control datasets. 84 (0.8%) of the 10,316 expressed repeat probes changed expression at least 2 fold ($p<0.01$) in the Dnmt TKO ES cells, with multiple probes for *MMERGLN* and *RLTR1B* (ERV1 family), and *IAP* and *RLTR45* (ERVK family) retrotransposons all showing upregulation in these cells (Figure 4A,

4B). These findings correlate well with recent RNA-seq data from Dnmt TKO ES cells: *MMERGLN*, *RLTR1B*, *IAP* and *RLTR45* are all upregulated ~2.5-13 fold in Dnmt TKO ES cell RNA-seq data [22]. However the two other elements (*MMERVK10C* and *RMER16*) reported as upregulated >2 fold in Dnmt TKO ES cells by RNA-seq (~2.3 fold upregulation for each [22]) have no detectable change in expression in the microarray data suggesting that microarray analysis is less sensitive than RNA-seq for detecting some changes in LTR retrotransposon expression. In *Eset*^{shRNA} ES cells, 125 (45%) of the 277 expressed repeat probes changed expression at least 2 fold ($p<0.01$), with multiple probes for *MMERGLN* (ERV1 family), *MMERVK10C, IAP* and *RLTR45* (ERVK family), *MERVL* (ERVL family) and *LINE-1* repetitive elements all showing upregulation in *Eset*^{shRNA} ES cells (Figure 4C, 4D). These elements represent a small subset of those reported previously as being uregulated in *Eset*$^{-/-}$ ES cells [22,24], which may reflect greater loss of *Eset* function in *Eset*$^{-/-}$ conditional knockout ES cells than in ES cells transiently transfected with knock-down shRNAs. Interestingly, although comparison of the Dnmt TKO and *Eset*^{shRNA} ES cell datasets suggest that some retrotransposon sequences (*MMERGLN*, *IAP*, *RLTR45*) are co-repressed by both DNA methyltransferases and Eset histone methyltransferase, analysis of the *Hdac1*$^{-/-}$ ES cell data shows striking divergences in the behaviour of these elements (Figure 4E, 4F). 74 (3.7%) of the 1971 expressed repeat probes changed expression at least 2 fold ($p<0.01$) in *Hdac1*$^{-/-}$ ES cells, with multiple probes for the *ETnERV3* and *RLTR45* (ERVK family) retrotransposons showing upregulation in *Hdac1*$^{-/-}$ ES cells (Figure 4E, 4F). These elements share considerable sequence similarity at the nucleotide level (84% identity over 4.2 kb of sequence). Interestingly, although *RLTR45* and *IAP* elements both appear to be co-repressed by DNA methyltransferases and Eset histone methyltransferase (Figure 4A-D), multiple probes for *IAP* (ERVK family) retrotransposons behaved quite differently from the *RLTR45* probes and were downregulated in *Hdac1*$^{-/-}$ ES cells (Figure 4E, 4F). Although Hdac1 typically acts as a transcriptional repressor, the apparent downregulation of *IAP* elements in *Hdac1*$^{-/-}$ ES cells would parallel the behaviour of some single-copy gene targets of Hdac1 [43]. We verified the microarray analysis of LTR retrotransposon

expression by performing qRT-PCR on *Hdac1*$^{-/-}$ ES cells: significant upregulation of *RLTR45* elements (11 fold, $p<0.05$) and downregulation of *IAP* elements (2.5 fold, $p<0.05$) was confirmed using this methodology (Figure 5A). Thus expression of some LTR retrotransposons is perturbed in the absence of *Hdac1* in mouse ES cells. Furthermore, the differences between *RLTR45* and *IAP* expression in *Hdac1*$^{-/-}$ ES cells suggests that an *Hdac1*-dependent transcriptional silencing mechanism is being recruited to retrotransposons independently of DNA methyltransferase or Eset histone methyltransferase activity.

The changes in *IAP* and *RLTR45* element expression in *Hdac1*$^{-/-}$ ES cells could be an indirect consequence of other gene expression changes that occur in *Hdac1*$^{-/-}$ ES cells [43], or may reflect a more direct role for Hdac1 in transcriptional regulation of these elements. To investigate whether *RLTR45* and *IAP* are direct targets of Hdac1 in ES cells, we analysed high throughput sequencing data from ES cell chromatin Hdac1 immunoprecipitation (Hdac1 ChIP-seq from mouse ES cells [38]) for enrichment of repetitive element sequences [25]. Interestingly, *RLTR45* LTR sequences are enriched in Hdac1 ChIP-seq relative to whole cell extract controls (Figure 5B), suggesting that Hdac1 is negatively regulating *RLTR45* expression in ES cells through physically associating with *RLTR45* LTRs. In contrast *IAP* LTR sequences are depleted in Hdac1 ChIP-seq (Figure 5B), consistent with the downregulation of *IAP* expression in *Hdac1*$^{-/-}$ ES cells being an indirect consequence of other changes in gene expression in these cells. Taken together, these data suggest that Hdac1 is directly recruited to *RLTR45* retrotransposons to silence their expression in ES cells.

Identifying LTR Retrotransposon Targets of Polycomb Repressive Complexes in ES Cells

Our genome-wide analysis of retrotransposon silencing in Dnmt TKO, *Eset*^{shRNA}, and *Hdac1*$^{-/-}$ ES cells suggests that multiple mechanisms contribute to silencing individual retrotransposon sequences in ES cells. These silencing mechanisms may be recruited sequentially or independently to target sequences. To investigate the interaction between different transcriptional repression

complexes at retrotransposon sequences in more detail, we examined retrotransposon silencing in ES cells carrying mutations in components of the polycomb repressive complexes PRC1 and PRC2. Conventional repression of gene expression by the polycomb repressive complexes PRC1 and PRC2 is thought to involve PRC2 methylating histone H3K27 and sequentially recruiting PRC1 to target loci [reviewed in 46]. However, a recent study on ES cells carrying mutations in the PRC1 component *Ring1B*, or mutations in the PRC2 component *Eed*, or mutations in both *Ring1B* and *Eed* has suggested that PRC1 and PRC2 are recruited independently and act redundantly to repress *MuLV* and *IAP* repetitive elements in this cell type [26]. We therefore computationally analysed genome-wide retrotransposon silencing in *Ring1B*$^{-/-}$, *Eed*$^{-/-}$, and *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ ES cells to determine whether any additional LTR retrotransposons are redundantly regulated by polycomb repressive complexes, and also to test whether any LTR retrotransposons are regulated by conventional sequential targeting of polycomb repressive complexes. *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ ES cells have numerous differences in gene expression compared to wild-type ES cells [26], and although LTR retrotransposon probes do not appear to be preferentially affected by loss of both PRC1 and PRC2 relative to other probes in the dataset, a number of ERV1 and ERVK probes are upregulated in *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ ES cells (Figure 6A). A smaller subset of LTR retrotransposon probes is upregulated in *Ring1B*$^{-/-}$ (Figure 6C) and *Eed*$^{-/-}$ (Figure 6E) single knockout ES cells. We identified LTR retrotransposon probes that were strongly upregulated at least 4 fold ($p<0.01$) in *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ ES cells (Figure 6B) and monitored how these LTR retrotransposons behaved in *Ring1B*$^{-/-}$ (Figure 6D) and *Eed*$^{-/-}$ (Figure 6F) single knockout ES cells. *MMVL30* (ERV1 family) probes were upregulated in *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ double knockout ES cells, but did not change greatly in either in *Ring1B*$^{-/-}$ or *Eed*$^{-/-}$ single knockout ES cells, consistent with these elements being redundantly and independently regulated by PRC1 and PRC2 [26]. A small number of *IAP* probes also appeared to be more strongly upregulated in *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ double knockout ES cells than in either single knockout cell line: 4 of the 112 *IAP* probes that are expressed in this dataset are upregulated at least 4 fold ($p<0.01$) in *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ double knockout ES cells, but no *IAP* probes are upregulated by these

criteria in either single knockout cell line (Figure 6B, 6D, 6F). This is consistent with previous observations that *IAP* elements are redundantly and independently regulated by PRC1 and PRC2 [26]. *RLTR45* (ERVK family) probes are also more strongly upregulated in *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ double knockout ES cells than in either single knockout cell line suggesting that this element is a novel retrotransposon target for redundant silencing by polycomb repressive complexes (Figure 6B, 6D, 6F).

Interestingly, genome-wide analysis of retrotransposon expression also suggests that some LTR retrotransposon probes are being repressed by conventional sequential recruitment of PRC2 and PRC1. *RLTR44* (ERVK family) probes appear to be similarly upregulated in *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ double knockout and single knockout ES cells (Figure 6B, 6D, 6F). The slightly lower upregulation of *RLTR44* probes in *Ring1B*$^{-/-}$ ES cells compared to *Eed*$^{-/-}$ ES cells may represent Ring1A-containing PRC1 complexes contributing to polycomb-mediated repression in ES cells [47]. *RLTR44* retrotransposons do however appear to be a novel retrotransposon target for conventional sequential silencing by polycomb repressive complexes. Thus computational analysis of gene expression in polycomb mutant cell lines suggests that PRC1 and PRC2 interact in different ways on different retrotransposon targets to bring about silencing of these repetitive elements in ES cells.

Differential Regulation of Retrotransposon Genomic Loci

During analysis of the *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ double knockout and single knockout ES cells, we noticed that probes for *RLTR4* retrotransposons were strongly upregulated in all three cell lines (Figure 6B, 6D, 6F). However the *RLTR4* probes that are upregulated correspond mainly to the LTR region (*RLTR4_Mm*) but usually not the internal region (*RLTR4-int*) of this element (Figure 7A). This suggests that the upregulation of these probes may represent expression from a subset of *RLTR4* loci, possibly corresponding to truncated or chimaeric elements. We therefore mapped the genomic location of the *RLTR4* LTR and internal probes that were upregulated in *Ring1B*$^{-/-}$ ES cells back

onto the genome using BLAT. In contrast to the retrotransposon probes upregulated in other datasets analysed in this study, the *RLTR4* probes upregulated in *Ring1B*$^{-/-}$ ES cells did not map to multiple genomic loci. Rather all of the upregulated *RLTR4* probes mapped only to a single *RLTR4*-containing genomic locus on chromosome 8 (chr8:125949704-125958431). The *RLTR4* probes that did not change expression in *Ring1B*$^{-/-}$ ES cells mapped to multiple loci in the genome. Thus the upregulation of a subset of *RLTR4* probes in *Ring1B*$^{-/-}$ ES cells may represent upregulation of a single genomic copy of this element. This locus appears to contain *RLTR4-int* and *MuLV-int* sequences flanked by *RLTR4_Mm* sequences that each contains an inversion and a ~200 bp deletion relative to the 742 bp consensus sequence. qRT-PCR using primers designed to specifically detect the *RLTR4-int* sequence at this locus confirmed that expression of this region is strongly upregulated in *Ring1B*$^{-/-}$ ES cells (Figure 7B), whereas qRT-PCR using primer sets that recognize multiple copies of *RLTR4-int* suggest that these elements are, in general, not upregulated in *Ring1B*$^{-/-}$ ES cells (Figure 7B). qRT-PCR also confirmed that representative ERV1, ERVK and ERVL LTR retrotransposons were not changing expression in *Ring1B*$^{-/-}$ ES cells (Figure 7B), consistent with the computational analysis. The divergent copy of *RLTR4* on chromosome 8 appears to be silenced by conventional polycomb repression as it is de-repressed in both *Ring1B*$^{-/-}$ and *Eed*$^{-/-}$ single knockout ES cells (Figure 6). This copy of *RLTR4* could have acquired *Ring1B* target sequences through mutations and re-arrangement to make it a target for conventional polycomb silencing. However as *RLTR4* is derived from *MuLV [48]*, a target of redundant silencing by PRC1 and PRC2 [26], it is perhaps more likely that changes in this divergent copy of *RLTR4* have removed sequences that allow PRC2-independent silencing of this locus by PRC1, making it behave as a conventional target for polycomb repression.

Many of the changes in retrotransposon expression that we have characterized in ES cells and germ cells involve subsets of probes for particular retrotransposons changing expression (Figure 1, 4, 5) suggesting that different genomic copies of these retrotransposons may be differentially regulated in

these cell types. In *Tex19.1*$^{-/-}$ testes, six of the nine expressed *MMERVK10C* probes in the dataset are upegulated at least 2 fold (Figure 1). All six of the upregulated *MMERVK10C* probes are located in the *MMERVK10C env* open reading frame. Two of the remaining three *MMERVK10C* probes are also located in the *env* gene and are upregulated in *Tex19.1*$^{-/-}$ testes, but are just below the 2 fold change threshold. The single *MMERVK10C* probe that is located in the *gag* region does not significantly change expression in the *Tex19.1*$^{-/-}$ testis dataset. We validated the computational data by qRT-PCR and confirmed that the *gag* and *env* regions of *MMERVK10C* are indeed differentially sensitive to loss of *Tex19.1* in mouse testes (Figure 7C). Interestingly, we noted that primer sets designed to different parts of *MMERVK10C env* (env.a – env.d) were also differentially sensitive to loss of *Tex19.1* (Figure 7C). These data suggest that a subset of *MMERVK10C* loci may be upregulated in *Tex19.1*$^{-/-}$ testes. Cloning and sequencing multiple independent clones of the env.c PCR product confirmed that multiple *MMERVK10C* loci were expressed in *Tex19.1*$^{-/-}$ and control testes (data not shown). The *pol* sequence is not covered by probes on the array but this region of *MMERVK10C* is also significantly upregulated in *Tex19.1*$^{-/-}$ testes (Figure 7C). Although *in silico* PCR suggests that the different *MMERVK10C* primer sets detect different numbers of *MMERVK10C* loci (*gag* primers detect 95 loci, *pol* primers detect 164 loci, *env.a – env.d* primers detect 78,70,179 and 40 loci respectively), the qRT-PCR data suggest that expression of these amplicons is differentially affected by loss of *Tex19.1*. We investigated the differential regulation of *MMERVK10C gag* and *env* regions by mapping the six strongly upregulated *env* probes and the single unaffected *gag* probe to individual *MMERVK10C* genomic loci, and assembled the *MMERVK10C* genomic loci into contigs. As *MMERVK10C* sequences that have retained flanking *RLTR10* LTRs are more likely to be transcriptionally active we selected *RLTR10*-flanked *MMERVK10C* contigs for further analysis (Figure 7D). Only 18 of the 250 *RLTR10*-flanked *MMERVK10C* contigs (7%) that we identified in the mouse genome are approximately full-length (contain >95% of *MMERVK10C* reference sequence). Interestingly, many of the *RLTR10*-flanked *MMERVK10C* contigs contain recurrent deletions: one recurrent deletion in the upregulated

MMERVK10C contigs removes the start of the *gag* open reading frame (nucleotides 399-870 deleted in 33% of these contigs) and appears to be associated with recurrent deletions in *env* (nucleotides 5810-6646 deleted in 33% of all contigs, 5810-6651 deleted in 20% of contigs). The presence of recurrent deletions in the *MMERVK10C* open reading frames at distinct genomic loci suggests that transcripts carrying these deletions may be actively retrotransposing, presumably in a non-autonomous manner through the activity of endogenous retroviral proteins provided *in trans*. The upregulated probes appeared to be highly representative of the *RLTR10*-flanked *MMERVK10C* loci, with 197 of the 250 *RLTR10*-flanked *MMERVK10C* contigs matching only the upregulated probes (Figure 7D). No *RLTR10*-flanked *MMERVK10C* contig matched all upregulated probes, or all the upregulated qRT-PCR primer sets, suggesting that multiple genomic copies of *MMERVK10C* are upregulated in *Tex19.1*$^{-/-}$ testes. In contrast, only two *RLTR10*-flanked *MMERVK10C* contigs matched only the unaffected probe (Figure 7D). Interestingly, 12 of the 15 *RLTR10*-flanked *MMERVK10C* contigs that matched both sets of probes were approximately full-length sequences, whereas the contigs that matched only the upregulated probes usually contained deletions with recurrent breakpoints. (Figure 7D). Furthermore, qRTPCR primers designed to amplify sequences within the 5810-6646 deletion (env.a) do not change expression in *Tex19.1*$^{-/-}$ testes, but those amplifying *env* sequences outside this deletion (env.b, env.c, and env.d) are upregulated (Figure 7C, 7D). Thus de-repression of specific subsets of MMERVK10C loci could be contributing to the differential regulation of different regions of *MMERVK10C gag* and *env* amplicons in *Tex19.1*$^{-/-}$ testes (Figure 7C). The upregulated pol and env.b/env.c primer sets can detect expression from *RLTR10*-flanked *MMERVK10C* contigs encoding intact pol and env proteins respectively (>90% of open reading frame intact relative to *MMERVK10C* reference sequence), but not contigs where the gag, pol, pro and env proteins are all intact. This suggests that the upregulated *MMERVK10C* transcripts may have some protein coding potential, but may need to rely on proteins provided *in trans* for retrotransposition. Some of the deletions in the upregulated *MMERVK10C* contigs, particularly the consistent disruption to parts of the *gag* region, may be removing sequences

used to recruit *Tex19.1*-independent retrotransposon silencing mechanisms. These loci would therefore be more reliant on the *Tex19.1*-dependent pathway for repression in germ cells, and be specifically de-repressed in *Tex19.1*$^{-/-}$ testes. Thus the differential regulation of *MMERVK10C* probes in *Tex19.1*$^{-/-}$ testes may be caused by the emergence of variant non-autonomous *MMERVK10C* elements that have deleted the sequences used to target silencing mechanisms to *MMERVK10C*.

We noted that *IAP* retrotransposon probes in Dnmt TKO ES cells lines were also exhibiting bimodal behaviour (Figure 4B). To investigate whether this represents differential regulation of *IAP* loci we designed qRT-PCR primers to *IAP* loci matching either upregulated or unaffected *IAP-int* probes (Figure 7E). qRT-PCR confirmed that some *IAP* loci are upregulated in Dnmt TKO ES cells, whereas others do not change expression (Figure 7F). As expected from the computational analysis of retrotransposon expression in Dnmt TKO ES cells, expression of *LINE-1* elements do not change in Dnmt TKO ES cells, and *MMERGLN* elements are upregulated, when assessed experimentally by qRT-PCR (Figure 7F). Our finding that different genomic copies of *IAP* may be differentially sensitive to loss of DNA methyltransferases is consistent with recent findings from RNA-seq of Dnmt TKO ES cells [22]. A simple interpretation of this phenomenon would be that the *IAP* loci that are not changing expression in Dnmt TKO ES cells are divergent defective copies of the *IAP* element. However, the unaffected *IAP-int* probes are detecting some *IAP* expression in ES cells, albeit at a lower level than the upregulated probes, suggesting that the *IAP* loci that are detected by the unaffected *IAP-int* probes are not all transcriptionally inert. To investigate why some *IAP* loci are insensitive to DNA methylation we identified the genomic *IAPEz-int* contigs that matched either the upregulated or the unaffected *IAP-int* probes. Although many of the contigs that only matched the unaffected *IAP-int* probes carried large deletions, one locus (chr10:22250294-22243066) contained a relatively intact *IAPEz-int* region flanked by *IAP* LTRs. Interestingly both of the LTRs at this locus contain a small 10 bp deletion (Figure 7G) that removes the conserved AP-1

transcription factor binding site [49]. Only 5 of the 16141 *IAP* LTRs in the mouse genome carry this, or a similar, deletion of the AP-1 binding site, and none of the *IAP* contigs that only match upregulated *IAP-int* probes contain this deletion in their LTRs. We confirmed that this copy of *IAP* (IAP_chr10) was not upregulated in Dnmt TKO ES cells by qRT-PCR (Figure 7F). However, mRNA from this locus was readily detected in wild-type and Dnmt TKO ES cells, suggesting that this copy of *IAP* is constitutively expressed. Loss of the AP-1 binding site in the *IAP* LTRs at this locus therefore does not appear to silence expression of this element, but may render this locus insensitive to regulation by DNA methylation. Interestingly, DNA methylation has been shown to inhibit binding of AP-1 to gene promoters [50]. Inhibition of AP-1 binding to *IAP* LTRs may therefore be contributing to DNA methylation-mediated repression of *IAP* elements in mouse ES cells.

Taken together, computational analysis of genome-wide retrotransposon silencing suggests that individual loci for a particular retrotransposon can have different sensitivities to retrotransposon suppression mechanisms. Mapping the changes that are present in differentially regulated loci may help to identify *cis*-acting retrotransposon sequences that are being used to recruit silencing mechanisms.

Discussion

Evaluation of the Microarray Repeat-Annotation Approach

In this manuscript we describe a simple computational approach to monitor repetitive element expression in microarray gene expression data. We have used repeat-annotation of pre-existing datasets to identify retrotransposons regulated by DNA methylation and different histone modifications in mouse ES cells (Table 3). We have verified that repeat probes present in gene expression microarrays are accurately reporting repetitive element expression by confirming our findings from *Tex19.1*, *Ring1B* and Dnmt TKO microarray analyses by qRT-PCR. In general there appears to be good qualitative correlation between repeats that we identified as changing expression in microarray datasets, and our qRT-PCR verification. Importantly there is also good correlation between repeat probes that are not changing expression in the microarray datasets and our qRT-PCR verification of these repetitive elements. Furthermore, we have used this approach to identify Hdac1 as a component of the retrotransposon silencing machinery in mouse ES cells (Figure 3, Figure 4). Application of this methodology to gene expression microarray data is likely to generate new insights into retrotransposon regulation in mammals, and help to identify further components of the defence mechanisms that protect the mammalian genome from retrotransposition. Consistent with previous re-annotation workflows designed to remove non-informative probes from microarray analyses [40], we found that commercially available mouse gene expression microarray platforms contain a number of probes that map to repetitive regions of the genome. Although expression information from these probes can be discarded to improve analysis of gene expression in the remaining dataset [40], we show here that the information from these probes can be extracted to accurately monitor repetitive element expression.

Repeat-annotation of microarray data can significantly expand the repertoire of repetitive elements studied in an experiment compared to testing selected representative candidates. Indeed this study has identified new target retrotransposons for polycomb repressive complexes and Hdac1 histone

deacetylase in mouse ES cells. Although the range of repetitive elements analysed by microarray repeat-annotation will not be as wide as that analysed by RNA-seq [22,23], between one and two thirds of all retrotransposons in the mouse genome are represented by probes on the microarray platforms that we have analysed here. A direct comparison between microarray repeat annotation (this study) and RNA-seq [22] for detecting changes in retrotransposon expression in Dnmt TKO ES cells shows good correlation between these methods (the four retrotransposons detected as upregulated by microarray analysis are the four most strongly upregulated retrotransposons detected by RNA-seq). However, two additional LTR retrotransposons were detected as upregulated in Dnmt TKO ES cells only by RNA-seq, despite representation of these elements on the microarray. Thus microarray analysis may be less sensitive than RNAseq for detecting some changes in LTR retrotransposon expression, particularly when only a small number of genomic copies are changing expression [22]. In addition, although we have focused on retrotransposon silencing in mouse germ cells and pluripotent cells, the computational approach that we describe here can be readily applied to microarray data from human cells and tissues to inform on retrotransposon expression in relation to retrotransposition in somatic mosaicism [12,37], epigenetic changes in cancer [51,52], reprogramming somatic cells into iPS cells [53], and toxicological insults [54]. As repeat-annotation can be applied to pre-existing microarray data as well as new datasets, this methodology can be used to extract information from many of the ~18,000 microarray gene expression data series that have been generated and deposited in publicly available databases [38]. This makes repeat-annotation of microarray data an attractive approach to test hypotheses and generate initial findings upon which more detailed research can be built. Thus microarray repeat-annotation represents a simple and cost-effective addition to the methods available to study repetitive element silencing at a genome-wide level.

Differential Regulation of Specific Genomic Copies of a Retrotransposon

One of the features of the computational approach that we have outlined here is that our analysis is based on aligning probe sequences to Repeatmasked regions of the genome, rather than to Repeatmasker consensus sequences. If different genomic copies of a repetitive element are behaving in different ways in an experiment then repeat-annotation of microarray data can potentially monitor expression from divergent genomic copies of a repetitive element. Clearly the extent to which multiple genomic copies of a particular element can be monitored will depend on the coverage of probes for that element. In the Affymetrix Mouse Expression 430 2.0 GeneChip platform that contains ~26,000 repeat probes we have been able to detect differential regulation of different genomic copies of *RLTR4* elements in *Ring1B*$^{-/-}$ ES cells, *IAP, RLTR45* and *RLTR1B* elements in Dnmt TKO ES cells and *ETnERV3* and *RLTR45* in *Hdac1*$^{-/-}$ ES cells. Remarkably, for *Ring1B*$^{-/-}$ ES cells we were able to detect expression changes that are possibly arising from only a single divergent copy of *RLTR4*. Thus repeat-annotation of microarray data appears to be able to monitor expression from divergent genomic copies of a repetitive element.

For *MMERVK10C* elements, analysis of the genomic loci matching retrotransposon probes was able to generate some insight into why some genomic copies of these elements are more sensitive to loss of suppression mechanisms than others. Loss of parts of the *gag* or *env* regions of *MMERVK10C* may be associated with genomic copies becoming more sensitive to *Tex19.1*-dependent suppression in male germ cells (Figure 7D). Interestingly, non-autonomous variants of *IAP* (*IAPΔ1*) that carry deletions in the *gag* region retrotranspose more frequently than their full-length counterparts [55]. Thus sequences in the *gag* region of both *IAP* and *MMERVK10* may be being used by host defence mechanisms to target these elements for silencing. In addition, analysis of differentially regulated *IAP* loci allowed us to identify a region in the *IAP* LTR that may be targeted by host silencing mechanisms (Figure 7G). DNA methylation at this conserved AP-1 transcription factor binding site may contribute to Dnmt-dependent repression of *IAP* elements in ES cells by inhibiting AP-1 binding. However, further experimental work is needed to functionally characterize the

consequences of these deletions for *MMERVK10C* and *IAP* silencing in germ cells and ES cells. Our analysis of *MMERVK10C* and *IAP* elements suggests that the behaviour of sequence variants in a retrotransposon's population can potentially be used to identify *cis*-acting sequences involved in retrotransposon suppression. In this respect, although repeat-annotation of microarray data may give some indication of differential regulation of repeat loci, RNA-seq may potentially be a more powerful approach to identify which genomic copies of an element are responsible for changes in expression.

As with all studies reporting changes in retrotransposon expression, determining whether changes in RNA or protein levels are caused by misregulation of one copy or many copies of a retrotransposon can be difficult. However, determining the sequence of the retrotransposon loci or transcripts that change expression in microarray datasets is an important prerequisite for assessing the functional potential of the mis-expressed retrotransposons. Finer sub-classification of repeat probes to distinguish between expression of functional and non-functional copies of a retrotransposon, for example active and inactive *LINE-1* elements, may not be accurate due to the short length of microarray probes: longer sequences are usually required to unambiguously identify a particular retrotransposon subfamily. Furthermore, none of the *LINE-1* probes present in the Illumina and Affymetrix arrays analysed here match the consensus monomer sequences that distinguish active Tf, Gf and A-type *LINE-1* elements. Thus microarray repeat-annotation may not be able to distinguish whether functional or non-functional genomic copies of a particular retrotransposon are deregulated, but may be useful in identifying subpopulations of genomic copies that include or exclude the misregulated retrotransposon sequence.

Regulation of Retrotransposon Expression in Mouse ES Cells and Germ Cells

We have used repeat-annotation of microarray data to investigate whether some of the established mechanisms for retrotransposon silencing have additional retrotransposon targets in mouse ES cells.

This analysis has demonstrated that there is a complex interplay between DNA methylation and histone modifications regulating the expression of the spectrum of repetitive elements in the mouse genome (Table 3). The LTR retrotransposons that we identified as being upregulated in Dnmt TKO ES cells overlap well with those identified recently by RNA-seq of Dnmt TKO ES cells [22]. Interestingly, many repetitive elements that belong to the same ERVK LTR retrotransposon family as *IAP* elements were not upregulated in Dnmt TKO ES cells suggesting that related retrotransposons can differ in their sensitivity to DNA methylation. Similarly, our finding that *MMERVK10C*, but not closely related retrotransposons such as *IAP*, are upregulated in *Tex19.1*$^{-/-}$ testes suggests that closely related retrotransposons differ in sensitivity to regulatory mechanisms in developing germ cells as well as ES cells. The differential behaviour of *IAP* and *MMERVK10C* elements in *Tex19.1*$^{-/-}$ testes could be caused by differences in the availability of transcriptional factors or by differences in silencing mechanisms associated with these elements. However as *IAP* LTRs are able to drive expression in spermatogonia [8,19], which are present in the 16 dpp *Tex19.1*$^{-/-}$ testes analysed here, the differential behaviour of *IAP* and *MMERVK10C* in *Tex19.1*$^{-/-}$ testes may reflect differences in silencing mechanisms acting on these elements. DNA methylation plays an important role in silencing *IAP* elements in spermatogonia [19], and redundancy between silencing mechanisms may well be contributing to the differential behaviour of *MMERVK10C* and *IAP* elements in *Tex19.1*$^{-/-}$ testes. Some of the retrotransposon targets for DNA methylation, *Tex19.1*, and the other silencing mechanisms that we have studied, may be obscured by redundancy between silencing mechanisms, and each of the mechanisms that we have studied here may have a broader range of targets than we have been able to identify.

Like *IAP* elements, the *RLTR45* ERVK LTR retrotransposon and the *MMERGLN* and *RLTR1B* ERV1 LTR retrotransposons are all upregulated in Dnmt TKO ES cells. The level of upregulation of *IAP*, *MMERGLN*, *RLTR45* and *RLTR1B* retrotransposons in Dnmt TKO ES cells was relatively low, consistent with previous observations for *IAP* elements [21]. Additional mechanisms are likely

to play a role in transcriptionally repressing these retrotransposons in ES cells, and Kap1/Eset-mediated repression appears to be one of the silencing pathways that plays a prominent role in repression of these elements [this study,22,23]. At least for *IAP* elements, differentiated cells may rely more heavily on DNA methylation than Kap1/Eset for repression [21,23]. It will be interesting to test whether transcription of *MMERGLN*, *RLTR1B* and *RLTR45* repetitive elements is directly regulated by DNA methylation, and whether DNA methylation plays a dominant role in repressing these repetitive elements in differentiated cells.

MMERGLN and *RLTR45* elements behaved similarly to *IAP* elements in $Eset^{shRNA}$ ES cells. Our finding that *MMERGLN*, *RLTR45* and *MMERVK10C* are all upregulated in $Eset^{shRNA}$ ES cells is consistent with these elements being enriched for H3K9Me3 in ES cells [22,25], and with recent RNA-seq and ChIP-seq data from $Eset^{-/-}$ ES cells [22]. We also found that *MERVL-int* elements were upregulated in $Eset^{shRNA}$ ES cells. These elements have also been reported to be upregulated in $Kap1^{-/-}$ ES cells [23]. ERVL retrotransposons are enriched for H3K27Me3 but not H3K9Me3 in ES cells [25], and the upregulation of *MERVL-int* (this study) and *MTA* [24] ERVL retrotransposons in $Eset^{shRNA}$ and $Eset^{-/-}$ ES cells may be an indirect effect of loss of *Eset* function. As ES cells lacking *Eset* differentiate towards the trophectoderm lineage [42], some of the changes in gene expression in $Eset^{shRNA}$ and $Eset^{-/-}$ ES cells may be an indirect consequence of this change in cell fate, or indeed any other change in gene expression. Indeed all of the microarray analyses of gene expression in ES cells that we have repeat-annotated are subject to the caveat that some changes in gene expression in these datasets may be consequences of differences in the proportion or type of differentiated cells present in the ES cell cultures. Further experiments will be required to determine why some ERVL retrotransposons are modestly upregulated in $Eset^{shRNA}$ and $Eset^{-/-}$ ES cells.

Importantly this study also identifies the histone deacetylase Hdac1 as a regulator of retrotransposon expression in mouse ES cells. The HDAC family of histone deacetylases has been

implicated in retrotransposon suppression in some cell types [28,44,45], and HDAC1 has been shown to suppress expression from avian retroviral LTRs in somatic HeLa cells [29,30]. The microarray analysis that we present here extends these findings by identifying the retrotransposon elements that are regulated by Hdac1 in mouse ES cells. Interestingly, although *RLTR45* and *IAP* elements behaved similarly in Dnmt TKO and *Eset*^{shRNA} ES cells, these elements were misregulated in opposite directions in *Hdac1*$^{-/-}$ ES cells. Thus the silencing mechanisms operating on repetitive elements appear to be modular, with different combinations of mechanisms acting on different elements (Table 3). Furthermore, these data suggest that the *Hdac1*-mediated and Dnmt-mediated silencing mechanisms operating on these elements are being targeted independently to *RLTR45* and *IAP* retrotransposons. The upregulation of *RLTR45* elements in *Hdac1*$^{-/-}$ ES cells, together with the enrichment of *RLTR45* sequences in Hdac1 ChIP-seq data from ES cells, suggests that an Hdac1-containing repressor complex may be recruited to *RLTR45* loci and silence this element. Further analysis of Hdac1-binding and histone modification at *RLTR45* elements is likely to generate more mechanistic insight into this silencing event. The downregulation of *IAP* elements in *Hdac1*$^{-/-}$ ES cells parallels the behaviour of some endogenous genes in these ES cells [43]. It will be informative to determine whether Hdac1 is acting directly on *IAP* elements to promote their transcriptional activation, or the increased activity of Hdac2 in *Hdac1*$^{-/-}$ ES cells is responsible for downregulation of *IAP* elements [43]. Interestingly, *LINE-1* elements did not appear to be upregulated in *Hdac1*$^{-/-}$ ES cells (Figure 5A), which contrasts with Hdac1's role in repressing *LINE-1* elements in neural stem cells [56]. Again, further experiments will be required to distinguish whether this difference reflects different chromatin environments between pluripotent ES cells and somatic neural stem cells, an effect of different Sox2-interacting partners in these cell types, or redundancy between multiple pathways operating to suppress *LINE-1* activity in ES cells.

In summary we have shown that genome-wide silencing of repetitive elements can be monitored by extracting this information from microarray gene expression data, revealing a complex interplay

between mechanisms that act to control retrotransposon expression in mouse ES cells and germ cells, and important differences in the behaviour of different genomic copies of individual retrotransposons. This computational approach has expanded our knowledge of retrotransposon targets for known silencing mechanisms, identified Hdac1 as a regulator of retrotransposon expression in ES cells, and demonstrated that epigenetic silencing mechanisms are independently recruited to retrotransposons in a modular manner.

Materials and Methods

Ethics Statement

Animal work was conducted according to UK Home Office regulations and local guidelines for animal welfare.

Animals

Mice were housed and bred according to UK Home Office regulations and local guidelines for animal welfare. *Tex19.1*$^{-/-}$ mice [39] were backcrossed three times to inbred C57BL/6 mice to reduce genetic variation prior to microarray analysis. Animals were culled at 16 days post partum (dpp) by cervical dislocation and testes from *Tex19.1*$^{-/-}$ experimental mice and *Tex19.1*$^{+/+}$ and *Tex19.1*$^{+/-}$ control littermates and frozen on liquid nitrogen prior to RNA isolation.

ES cell culture

Ring1B$^{-/-}$ feeder-dependent ES cells [57] were cultured at 37°C, 5% CO_2 on a mitomycin C-treated mouse embryonic fibroblast feeder layer, feeder-independent Dnmt TKO and *Hdac1*$^{-/-}$ ES cells [20,43]were cultured at 37°C, 5% CO_2 on gelatinized tissue-culture flasks. *Ring1B*$^{-/-}$ and *Hdac1*$^{-/-}$ ES cells were cultured using DMEM (Invitrogen) supplemented with 15% fetal calf serum, 1 x non-essential amino acids, 1 mM sodium pyruvate, 100 units/mL penicillin, 0.1 mg/mL streptomycin, 50 μM β-mercaptethanol and leukemia inhibitory factor. Dnmt TKO ES cells were cultured using GMEM (Invitrogen) with 10% fetal calf serum rather than DMEM with 15% fetal calf serum. ES cells were harvested using trypsin-EDTA, then pelleted for RNA isolation.

RNA isolation and Illumina Beadarray Gene Expression Profiling

RNA was isolated from 16 dpp testis or ES cell pellets using TRIzol (Invitrogen) according to the manufacturer's instructions. RNA was treated with DNAseI (Roche) for 2h at 37°C to remove genomic DNA contamination. For Illumina Beadarrays of 16 dpp testes, cRNA samples were

prepared using Illumina TotalPrep RNA Amplification Kit (Ambion) and hybridized to Illumina Mouse WG-6 v2.0 Beadarrays according to the manufacturers' protocols. The raw and processed *Tex19.1* microarray data have been deposited in the publicly accessible GEO database [38], accesssion number GSE30461.

qRT-PCR

cDNA synthesis was performed on DNaseI-treated RNA using random primers and Superscript III reverse transcriptase (Invitrogen) according to the manufacturer's instructions. qRT-PCR was performed using Brilliant II/III SYBR Green QPCR Master Mix (Agilent Technologies) or Quantitect SYBR Green detection kit (Qiagen) and a CFX96 Real-Time PCR Detection System (Bio-Rad). Relative changes in gene expression were calculated by normalising gene expression levels from different samples to *β-Actin* or *Gapdh* as indicated. Expression levels in experimental samples are expressed relative to wild type ES cells or littermate controls. Three technical replicates were performed for each biological sample, and cDNA prepared from each RNA sample in the absence of reverse transcriptase showed no significant qRT-PCR signals. For qRT-PCR of 16 dpp testes, the Sertoli cell-expressed *Sdmg1* gene [58] was used to verify normalization between animals. A two-tailed t-test was used to determine statistical significance of qRT-PCR gene expression changes. The sequences of the primers used for qRT-PCR are available with the online version of this paper (Supporting Figure S7).

Repeat Annotation of Illumina Probes

The DNA sequence of the 50-mer probes used in Illumina Mouse WG-6 Beadchips were downloaded from the manufacturers website [59]. For each Beadchip version the probe sequences were used to search the mm9 release of the mouse genome by individual chromosome using BLAT [60]. The BLAT parameters used were -minIdentity=95 -stepSize=5 -repMatch=2253. Experimental data suggests that the 50 nt Illumina Beadchip probes will hybridize to mRNAs containing 2

mismatches in the probe sequence with an efficiency of greater than 90% [61]. A *Perl* script was used to compare the genome co-ordinates of the top hit (with a 48/50 nt identity minimum cut-off) for each probe sequence with the co-ordinates of the Repeatmasked regions of the mm9 release of the mouse genome downloaded from the UCSC genome browser [62]. Probes that overlapped a Repeatmasked region by at least 48 nt, and were in the appropriate orientation to recognize sense transcripts were selected and annotated with the Repeatmasker class, family and element corresponding to that genomic region. Tables containing the repetitive element probes for each Illumina Beadchip were imported into R [63] and used to annotate Illumina Beadchip gene expression data. These annotation tables are available with the online version of this paper (Supporting Dataset S2-S4).

Pre-processing of Illumina Beadchip Gene Expression Data

Illumina Beadchip gene expression data for *Eset* shRNA knock-down ES cells were downloaded from the NCBI GEO repository [38], accession number GSE17439 [42]. All analysis of Illumina Mouse Whole Genome WG-6 Beadchip microarrays was performed on probe-level data. Probe-level expression data were background-subtracted in Illumina Beadstudio, then imported into the *lumi* Bioconductor package [64] in R. The data were then log-transformed, and quantile-normalized in *lumi*. The expression data and present/absent calls were exported from *lumi* and any probes that were called as absent in all samples in the experiment were removed from the dataset.

Repeat Annotation of Affymetrix Probes

These Affymetrix Murine Genome U74Av2 and Mouse Expression 430 2.0 GeneChips contain ~12,000 and ~45,000 probesets respectively, with each probeset containing ~11 different 25 nt probes targeting a specific transcript. The DNA sequence of the 25-mer probes used in Affymetrix Mouse Gene Expresson Arrays were downloaded from the manufacturer's website [65]. For each version of these arrays the probe sequences were used to BLAT search the mm9 release of the

mouse genome by individual chromosome. The BLAT search parameters were -minIdentity=95 -tileSize=11 -stepSize=5 -repMatch=1[10]. The genome co-ordinates of the top hit (with a 24/25 nt identity minimum cut-off) for each probe sequence were compared to the co-ordinates of the Repeatmasked regions of the mm9 release of mouse genome using a *Perl* script. Probes that overlapped a Repeatmasked region by at least 24 nt, and were in the appropriate orientation to recognize sense transcripts, were selected and annotated with the Repeatmasker class, family, and element corresponding to that genomic region. Tables containing the repetitive element probes for each Affymetrix array platform were imported into R. These annotation tables are available with the online version of this paper (Supporting Dataset S5, S6).

Pre-processing of Affymetrix Microarray Gene Expression Data

Affymetrix Mouse Gene Expression data for *Ring1B*$^{-/-}$, *Eed*$^{-/-}$, *Ring1B*$^{-/-}$ *Eed*$^{-/-}$, *Hdac1*$^{-/-}$ and Dnmt TKO ES cells were downloaded from the NCBI GEO repository [38], accession numbers GSE19076 [26], GSE20177 [41] and GSE5583 [43]. Raw Affymetrix data were imported into the *affy* Bioconductor package [66] in R. Probe expression values were background-corrected using the robust multi-array average algorithm [67] in *affy*. Expression values for the perfect match probes were extracted from *affy*, log-transformed, then quantile-normalized. Summation across probesets was not performed so that the Affymetrix data could be analysed at the probe level. Probes that were expressed at more than the sample median level in at least half the arrays for one experimental condition in a dataset were considered to be present [68]. Absent probes were removed from the dataset to simplify the analysis. Some probe sequences in the Affymetrix Gene Expression platform are present in more than one probeset, and these redundant probes are present at multiple locations in the array. Therefore some 25-mer DNA sequences are represented by more than one probe in the Affymetrix datasets.

Identification of Differentially Expressed Probes in Illumina and Affymetrix Microarray Data

For both Illumina and Affymetrix data, the R Bioconductor package *limma* [69] was used to identify probes that were expressed at different levels in experimental and control conditions by linear modeling. The Benjamini-Hochberg method was used to correct for multiple testing in *limma*, and adjusted p-values of ≤0.01 were considered to be statistically significant. Tables corresponding to all expressed probes in the experiment, and probes that statistically changed during the experiment, were repeat-annotated in R using the tables generated in sections 2.5 and 2.7. The resulting data were graphed using R. MA-plots show the log fold change in expression of each probe, plotted against the average expression of that probe in the dataset. Probability density functions for the microarray data were generated by kernel density estimation in R.

Phylogenetic Analysis

Close relatives of *MMERVK10C* were found by using *MMERVK10C* as a template for Genewise [70] to predict *pol* and *pro* sequences in the Repbase database of repetitive DNA sequences [3]. Multiple protein alignment was performed using ClustalW [71], and phylogenetic trees were constructed using MEGA4 [72] to apply the neighbour-joining method [73]. Phylogenies were based on the proportion of amino acid sites at which sequences are different, with pairwise deletion to remove gaps in alignments as the need arises. The reliability of each interior branch of a given topology was assessed using the bootstrap interior branch test with 1000 bootstraps.

Assembly of Repeatmasker Genomic Hits into Contigs

The co-ordinates of the Repeatmasked regions of the mm9 release of the mouse genome were downloaded from the UCSC genome browser [62], and regions Repeatmasked for *MMERVK10C-int* or *IAP-int* were extracted. The hits were ordered by their co-ordinates and adjacent hits that were in the same orientation on the same chromosome, were collinear on the consensus sequence, and were separated by less than the length of the consensus sequence were assembled into the same contig. *IAP*-int contigs that had *IAP* LTRs located within 50 bp of both ends of the contig were

identified for further analysis. A similar approach was used to identify *RLTR10*-flanked *MMERVK10C* contigs, with *RLTR10* genomic loci greater than 250 bp included in the assembly.

LTR Retrotransposon Enrichment in ChIP-seq Data

Hdac1 ES cell ChIP-seq and control ES cell whole cell extract datasets were downloaded from the GEO repository (accession number GSE27844). LTR retrotransposon enrichment was calculated using the Repeat Enrichment Estimator web application [25], and data for either all *IAP* LTR sequences, or all *RLTR45* LTR sequences, were combined.

Authors' contributions

JR performed the mouse breeding, the *Tex19.1* microarray, and helped draft the manuscript. JR, JC and MJM performed the qRT-PCR verification of the microarray repeat-annotations. JC performed computational analysis of Hdac1 ChIP-seq data, MT cloned and sequenced *MMERVK10C* RT-PCR products. PG performed phylogenetic analysis of *MMERVK10C* elements and helped in the design of the study. JLG-P and RRM participated in the organization and design of the study and helped draft the manuscript. IRA developed the repeat-annotation methodology, performed the microarray repeat-annotations, co-ordinated the study and drafted the manuscript.

Acknowledgements

We thank Masaki Okano (Kobe, Japan), Christian Seiser (Vienna, Austria), and Anton Wutz (Cambridge, UK) for Dnmt TKO, *Hdac1*$^{-/-}$, and *Ring1B*$^{-/-}$ ES cells respectively. We thank the Wellcome Trust Clinical Research Facility at the Western General Hospital in Edinburgh for performing the Illumina Beadarray hybridizations, and Wendy Bickmore and Sara Heras (both MRC-HGU, Edinburgh) for helpful suggestions and critical reading of the manuscript.

References

1. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921. doi:10.1038/35057062.

2. International Mouse Genome Sequencing Consortium (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420: 520–562. doi:10.1038/nature01262.

3. Jurka J (1998) Repeats in genomic DNA: mining and meaning. Curr Opin Struct Biol 8: 333–337.

4. Kleckner N (1990) Regulation of Transposition in Bacteria. Annu Rev Cell Biol 6: 297–327. doi:10.1146/annurev.cb.06.110190.001501.

5. Garfinkel DJ, Boeke JD, Fink GR (1985) Ty element transposition: Reverse transcriptase and virus-like particles. Cell 42: 507–517. doi:10.1016/0092-8674(85)90108-4.

6. Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: A mechanism for non-LTR retrotransposition. Cell 72: 595–605. doi:10.1016/0092-8674(93)90078-5.

7. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, et al. (1996) High Frequency Retrotransposition in Cultured Mammalian Cells. Cell 87: 917–927. doi:10.1016/S0092-8674(00)81998-4.

8. Dewannieux M, Heidmann T (2005) L1-mediated Retrotransposition of Murine B1 and B2 SINEs Recapitulated in Cultured Cells. Journal of Molecular Biology 349: 241–247. doi:16/j.jmb.2005.03.068.

9. Heidmann O, Heidmann T (1991) Retrotransposition of a mouse IAP sequence tagged with an indicator gene. Cell 64: 159–170. doi:10.1016/0092-8674(91)90217-M.

10. Goodier JL, Kazazian HH (2008) Retrotransposons Revisited: The Restraint and Rehabilitation of Parasites. Cell 135: 23–35. doi:10.1016/j.cell.2008.09.022.

11. Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, et al. (2011) Somatic retrotransposition alters the genetic landscape of the human brain. Nature. doi:10.1038/nature10531.

12. Muotri AR, Chu VT, Marchetto MCN, Deng W, Moran JV, et al. (2005) Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature 435: 903–910. doi:10.1038/nature03663.

13. Smit AFA, Hubley R, Green P (2011) Repeatmasker - Home. Available:http://www.repeatmasker.org/. Accessed 12 June 2011.

14. Maksakova IA, Romanish MT, Gagnier L, Dunn CA, van de Lagemaat LN, et al. (2006) Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. PLoS Genet 2: e2. doi:10.1371/journal.pgen.0020002.

15. Zamudio N, Bourc'his D (2010) Transposable elements in the mammalian germline: a comfortable niche or a deadly trap? Heredity 105: 92–104. doi:10.1038/hdy.2010.53.

16. Rowe HM, Trono D (2011) Dynamic control of endogenous retroviruses during development. Virology 411: 273–287. doi:10.1016/j.virol.2010.12.007.

17. Öllinger R, Reichmann J, Adams IR (2010) Meiosis and retrotransposon silencing during germ cell development in mice. Differentiation 79: 147–158. doi:10.1016/j.diff.2009.10.004.

18. Walsh CP, Chaillet JR, Bestor TH (1998) Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 20: 116–117. doi:10.1038/2413.

19. Bourc'his D, Bestor TH (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 431: 96–99. doi:10.1038/nature02886.

20. Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S-ichiro, Sakaue M, et al. (2006) Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells 11: 805–814. doi:10.1111/j.1365-2443.2006.00984.x.

21. Hutnick LK, Huang X, Loo T-C, Ma Z, Fan G (2010) Repression of Retrotransposal Elements in Mouse Embryonic Stem Cells Is Primarily Mediated by a DNA Methylation-independent Mechanism. Journal of Biological Chemistry 285: 21082–21091. doi:10.1074/jbc.M110.125674.

22. Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, et al. (2011) DNA Methylation and SETDB1/H3K9me3 Regulate Predominantly Distinct Sets of Genes, Retroelements, and Chimeric Transcripts in mESCs. Cell Stem Cell 8: 676–687. doi:10.1016/j.stem.2011.04.004.

23. Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, et al. (2010) KAP1 controls endogenous retroviruses in embryonic stem cells. Nature 463: 237–240. doi:10.1038/nature08674.

24. Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, et al. (2010) Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature 464: 927–931. doi:10.1038/nature08858.

25. Day DS, Luquette LJ, Park PJ, Kharchenko PV (2010) Estimating enrichment of repetitive elements from high-throughput sequence data. Genome Biol 11: R69–R69. doi:10.1186/gb-2010-11-6-r69.

26. Leeb M, Pasini D, Novatchkova M, Jaritz M, Helin K, et al. (2010) Polycomb complexes act redundantly to repress genomic repeats and genes. Genes & Development 24: 265–276. doi:10.1101/gad.544410.

27. Macfarlan TS, Gifford WD, Agarwal S, Driscoll S, Lettieri K, et al. (2011) Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. Genes & Development 25: 594–607. doi:10.1101/gad.2008511.

28. Garcia-Perez JL, Morell M, Scheys JO, Kulpa DA, Morell S, et al. (2010) Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. Nature 466: 769–773. doi:10.1038/nature09209.

29. Poleshko A, Einarson MB, Shalginskikh N, Zhang R, Adams PD, et al. (2010) Identification of a functional network of human epigenetic silencing factors. J Biol Chem 285: 422–433. doi:10.1074/jbc.M109.064667.

30. Poleshko A, Palagin I, Zhang R, Boimel P, Castagna C, et al. (2008) Identification of cellular proteins that maintain retroviral epigenetic silencing: evidence for an antiviral response. J Virol 82: 2313–2323. doi:10.1128/JVI.01882-07.

31. Murchison EP, Stein P, Xuan Z, Pan H, Zhang MQ, et al. (2007) Critical roles for Dicer in the female germline. Genes Dev 21: 682–693. doi:10.1101/gad.1521307.

32. Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, et al. (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature 453: 534–538. doi:10.1038/nature06904.

33. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, et al. (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453: 539–543. doi:10.1038/nature06908.

34. Holmes RK, Malim MH, Bishop KN (2007) APOBEC-mediated viral restriction: not simply editing? Trends in Biochemical Sciences 32: 118–128. doi:16/j.tibs.2007.01.004.

35. Stetson DB, Ko JS, Heidmann T, Medzhitov R (2008) Trex1 Prevents Cell-Intrinsic Initiation of Autoimmunity. Cell 134: 587–598. doi:10.1016/j.cell.2008.06.032.

36. Wissing S, Montano M, Garcia-Perez JL, Moran JV, Greene WC (2011) Endogenous APOBEC3B Restricts LINE-1 Retrotransposition in Transformed Cells and Human Embryonic Stem Cells. J Biol Chem 286: 36427–36437. doi:10.1074/jbc.M111.251058.

37. Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, et al. (2009) The regulated retrotransposon transcriptome of mammalian cells. Nat Genet 41: 563–571. doi:10.1038/ng.368.

38. Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, et al. (2010) NCBI GEO: archive for functional genomics data sets—10 years on. Nucleic Acids Research 39: D1005–D1010. doi:10.1093/nar/gkq1184.

39. Öllinger R, Childs AJ, Burgess HM, Speed RM, Lundegaard PR, et al. (2008) Deletion of the pluripotency-associated Tex19.1 gene causes activation of endogenous retroviruses and defective spermatogenesis in mice. PLoS Genet 4: e1000199. doi:10.1371/journal.pgen.1000199.

40. Barbosa-Morais NL, Dunning MJ, Samarajiwa SA, Darot JFJ, Ritchie ME, et al. (2010) A re-annotation pipeline for Illumina BeadArrays: improving the interpretation of gene expression data. Nucleic Acids Research 38: e17. doi:10.1093/nar/gkp942.

41. Sakaue M, Ohta H, Kumaki Y, Oda M, Sakaide Y, et al. (2010) DNA methylation is dispensable for the growth and survival of the extraembryonic lineages. Curr Biol 20: 1452–1457. doi:10.1016/j.cub.2010.06.050.

42. Yuan P, Han J, Guo G, Orlov YL, Huss M, et al. (2009) Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. Genes Dev 23: 2507–2520. doi:10.1101/gad.1831909.

43. Zupkovitz G, Tischler J, Posch M, Sadzak I, Ramsauer K, et al. (2006) Negative and positive regulation of gene expression by mouse histone deacetylase 1. Mol Cell Biol 26: 7913–7928. doi:10.1128/MCB.01220-06.

44. Yang X-J, Seto E (2008) The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. Nat Rev Mol Cell Biol 9: 206–218. doi:10.1038/nrm2346.

45. Brunmeir R, Lagger S, Simboeck E, Sawicka A, Egger G, et al. (2010) Epigenetic Regulation of a Murine Retrotransposon by a Dual Histone Modification Mark. PLoS Genet 6: e1000927. doi:10.1371/journal.pgen.1000927.

46. Simon JA, Kingston RE (2009) Mechanisms of Polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol 10: 697–708. doi:10.1038/nrm2763.

47. Endoh M, Endo TA, Endoh T, Fujimura Y-ichi, Ohara O, et al. (2008) Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. Development 135: 1513–1524. doi:10.1242/dev.014340.

48. Changolkar LN, Singh G, Pehrson JR (2008) macroH2A1-Dependent Silencing of Endogenous Murine Leukemia Viruses. Mol Cell Biol 28: 2059–2065. doi:10.1128/MCB.01362-07.

49. Falzon M, Kuff EL (1988) Multiple protein-binding sites in an intracisternal A particle long terminal repeat. J Virol 62: 4070–4077.

50. Dong Z, Wang X, Evers BM (2000) Site-specific DNA methylation contributes to neurotensin/neuromedin N expression in colon cancers. American Journal of Physiology - Gastrointestinal and Liver Physiology 279: G1139–G1147.

51. Goering W, Ribarska T, Schulz WA (2011) Selective changes of retroelement expression in human prostate cancer. Carcinogenesis 32: 1484–1492. doi:10.1093/carcin/bgr181.

52. Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, et al. (2011) Aberrant Overexpression of Satellite Repeats in Pancreatic and Other Epithelial Cancers. Science. doi:10.1126/science.1200801.

53. Wissing S, Muñoz-Lopez M, Macia A, Yang Z, Montano M, et al. (2011) Reprogramming somatic cells into iPS cells activates LINE-1 retroelement mobility. Human Molecular Genetics. doi:10.1093/hmg/ddr455.

54. Irons RD, Stillman WS, Cloyd MW (1987) Selective activation of endogenous ecotropic retrovirus in hematopoietic tissues of B6C3F1 mice during the preleukemic phase of 1,3-butadiene exposure. Virology 161: 457–462. doi:10.1016/0042-6822(87)90139-5.

55. Horie K, Saito E-suke, Keng VW, Ikeda R, Ishihara H, et al. (2007) Retrotransposons Influence the Mouse Transcriptome: Implication for the Divergence of Genetic Traits. Genetics 176: 815–827. doi:10.1534/genetics.107.071647.

56. Muotri AR, Marchetto MCN, Coufal NG, Oefner R, Yeo G, et al. (2010) L1 retrotransposition in neurons is modulated by MeCP2. Nature 468: 443–446. doi:10.1038/nature09544.

57. Leeb M, Wutz A (2007) Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. J Cell Biol 178: 219–229. doi:10.1083/jcb.200612127.

58. Best D, Sahlender DA, Walther N, Peden AA, Adams IR (2008) Sdmg1 is a conserved transmembrane protein associated with germ cell sex determination and germline-soma interactions in mice. Development 135: 1415–1425. doi:10.1242/dev.019497.

59. Illumina, Inc. (2011) Illumina Resources - Annotation Files. Available:http://www.switchtoi.com/annotationfiles.ilmn. Accessed 12 June 2011.

60. Kent WJ (2002) BLAT—The BLAST-Like Alignment Tool. Genome Research 12: 656–664. doi:10.1101/gr.229202.

61. He Z, Wu L, Li X, Fields MW, Zhou J (2005) Empirical Establishment of Oligonucleotide Probe Design Criteria. Appl Environ Microbiol 71: 3753–3760. doi:doi:10.1128/AEM.71.7.3753-3760.2005.

62. Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, et al. (2010) The UCSC Genome Browser database: update 2011. Nucleic Acids Research. doi:10.1093/nar/gkq963.

63. R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/. Available:http://www.r-project.org/. Accessed 12 June 2011.

64. Du P, Kibbe WA, Lin SM (2008) lumi: a pipeline for processing Illumina microarray. Bioinformatics 24: 1547–1548. doi:10.1093/bioinformatics/btn224.

65. Affymetrix, Inc. (2011) Affymetrix - Home. Available:http://www.affymetrix.com/. Accessed 12 June 2011.

66. Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20: 307–315. doi:10.1093/bioinformatics/btg405.

67. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4: 249–264. doi:10.1093/biostatistics/4.2.249.

68. Laajala E, Aittokallio T, Lahesmaa R, Elo LL (2009) Probe-level estimation improves the detection of differential splicing in Affymetrix exon array studies. Genome Biol 10: R77–R77. doi:10.1186/gb-2009-10-7-r77.

69. Smyth GK (2004) Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. Statistical Applications in Genetics and Molecular Biology 3. doi:10.2202/1544-6115.1027.

70. Birney E, Clamp M, Durbin R (2004) GeneWise and Genomewise. Genome Res 14: 988–995. doi:10.1101/gr.1865504.

71. Higgins DG, Thompson JD, Gibson TJ (1996) Using CLUSTAL for multiple sequence alignments. Meth Enzymol 266: 383–402.

72. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599. doi:10.1093/molbev/msm092.

73. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425.

Figure Legends

Figure 1. Genome-Wide Repetitive Element Expression in *Tex19.1*$^{-/-}$ Testes

(A-C) MA-plots showing the mean expression level for each expressed probe in the *Tex19.1* testis Illumina Beadarray data plotted against the fold upregulation of that probe in *Tex19.1*$^{-/-}$ testes. Probes for repeat families (A), classes of LTR retrotransposons (B), and the *MMERVK10C* element (C) are colour-coded in each plot according to the legend. Note the group of six *MMERVK10C* ERVK LTR retrotransposon probes upregulated in *Tex19.1*$^{-/-}$ testes. (D) Plot showing the behaviour of the entire *MMERVK10C* probe population in *Tex19.1*$^{-/-}$ testes. Vertical lines indicate a 2 fold change. (E) qRT-PCR verification of *MMERVK10C* upregulation in C57BL/6 *Tex19.1*$^{-/-}$ testes. Expression levels for each repetitive element (mean ± standard error for three animals) were normalized to *β-Actin* and expressed relative to littermate controls. Representative LTR retrotransposons belonging to ERV1, ERVK and ERVL classes do not change expression in *Tex19.1*$^{-/-}$ testes. *Sdmg1* is a single-copy control gene for Sertoli cell expression to verify normalization between animals. MMERVK10C env.c and LINE1 ORF2 primer sets (Supporting Figure S7) were used to assess *MMERVK10C* and *LINE-1* expression. Asterisk indicates a statistically significant difference ($p<0.05$).

Figure 2. Closely Related Retrotransposons are Differentially Sensitive to Loss of *Tex19.1*

(A) Phylogeny of mouse retrotransposon pol and pro proteins. *MMERVK10C* sequences are highlighted in red. The *MMERVK10C* sequences lie within a cluster of *IAP*-type sequences (yellow). (B) Plot showing the likelihood of *IAP* probes changing expression in the *Tex19.1*$^{-/-}$ microarray dataset. (C) qRT-PCR for retrotransposons closely related to *MMERVK10C* in *Tex19.1*$^{-/-}$ knockout and littermate control testes at 16 dpp. Expression levels for each repetitive element (mean ± standard error for three animals) were normalized to *β-Actin* and expressed relative to littermate controls. MMERVK10C env.c and IAP primer sets (Supporting Figure S7) were used to

assess *MMERVK10C* and *IAPEz* expression. Asterisk indicates a statistically significant difference ($p<0.05$).

Figure 3. Different Transcriptional Silencing Mechanisms Have Distinct Effects on Genome-Wide Repression of Repetitive Elements

(A, B) MA-plots for *Dnmt1*$^{-/-}$ *Dnmt3A*$^{-/-}$ *Dnmt3B*$^{-/-}$ triple knockout (Dnmt TKO) ES cell Affymetrix Gene Expression data. The mean expression level for each expressed probe is plotted against the fold upregulation of that probe in Dnmt TKO ES cells. Probes for repeat families (A), and classes of LTR retrotransposons (B) are colour-coded in each plot according to the legend. A group of ERV1 and ERVK LTR retrotransposons can be seen to be upregulated relative to the total probe population in the Dnmt TKO ES cells. (C, D) MA-plots for *Eset*^{shRNA} ES cell Illumina Beadchip data with probes for repeat families (C), and classes of LTR retrotransposons (D) colour-coded according to the legend. Probes for different ERV1, ERVK and ERVL LTR retrotransposon families are strongly upregulated, and multiple *LINE-1* probes are modestly upregulated, in *Eset*^{shRNA} ES cells. (E, F) MA-plots for *Hdac1*$^{-/-}$ ES cell Affymetrix Gene Expression data with probes for repeat families (E), and classes of LTR retrotransposons (F) colour-coded according to the legend. One group of ERVK LTR retrotransposon probes is upregulated in *Hdac1*$^{-/-}$ ES cells, another group is downregulated.

Figure 4. Genome-Wide Retrotransposon Targets of Transcriptional Repression Mechanisms in Mouse ES Cells

(A, C, E) Histograms showing repeat probes that change expression at least 2 fold ($p<0.01$) in Dnmt TKO, *Eset*^{shRNA}, and *Hdac1*$^{-/-}$ ES cells respectively. (B, D, F) Plots showing the behaviour of the selected retrotransposon probe populations in Dnmt TKO, *Eset*^{shRNA}, and *Hdac1*$^{-/-}$ ES cells respectively. Retrotransposons are colour-coded according to the legend. Vertical lines indicate the

2 fold change cut-off used in panels A, C and E. Note the divergent behaviour of *IAP* and *RLTR45* retrotransposons in *Hdac1*$^{-/-}$ ES cells in contrast to Dnmt TKO and *Eset*^{shRNA} ES cells.

Figure 5. Hdac1 Regulates Expression of LTR Retrotransposons in Mouse ES Cells

(A) qRT-PCR verification of *LINE-1*, *RLTR45* and *IAP* expression in *Hdac1*$^{-/-}$ ES cells. Expression levels (mean ± standard error for three biological replicates) were normalized to *β-Actin* and expressed relative to control ES cells. IAP and LINE1 5'UTR primer sets (Supplementary Figure S7) were used to assess *IAP* and *LINE-1* expression. Asterisks indicate a statistically significant difference ($p<0.05$) for *RLTR45* and *IAP* elements. *RLTR45* expression is upregulated in *Hdac1*$^{-/-}$ ES cells, but *IAP* expression is downregulated. (B) Enrichment of LTR retrotransposon sequences in Hdac1 ChIP-seq data from mouse ES cells. The maximum likelihood of enrichment (± 95% confidence intervals) for *RLTR45* LTR and *IAP* LTR sequences Hdac1 ChIP-seq relative to whole cell extract is shown. *RLTR45* LTR sequences are enriched in the Hdac1 ChIP-seq indicating a physical association between Hdac1 and *RLTR45* retrotransposon chromatin, in contrast *IAP* LTR sequences are depleted.

Figure 6. LTR Retrotransposon Targets of Polycomb Repressive Complexes in ES Cells

(A, C, E) MA-plots for *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ double knockout, *Ring1B*$^{-/-}$ single knockout and *Eed*$^{-/-}$ single knockout ES cells showing how different classes of LTR retrotransposons change expression in these cell lines. (B, D, F) Plots showing the behaviour of selected retrotransposon probe populations in *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ double knockout, *Ring1B*$^{-/-}$ single knockout and *Eed*$^{-/-}$ single knockout ES cells. The selected retrotransposons are all represented by multiple upregulated probes (≥4 fold upregulation, $p<0.01$) in *Ring1B*$^{-/-}$ *Eed*$^{-/-}$ ES cells. Vertical lines indicate a 4 fold change. Note that some retrotransposons (e.g. *MMVL30*, *RLTR45*) are upregulated in double knockout but not single knockout ES cells, other retrotransposons (e.g. *RLTR44*) are upregulated in all three ES cell lines. Retrotransposon probes are colour-coded as shown in the plot legends.

Figure 7. Differential Regulation of Retrotransposon Genomic Loci

(A) Plot showing the differential behaviour of different *RLTR4* retrotransposon probe populations in *Ring1B*$^{-/-}$ single knockout ES cells. Different *RLTR4* probe populations are colour-coded as shown in the legend, and vertical lines indicate a 4 fold change. (B) qRT-PCR verification of repetitive element expression in *Ring1B*$^{-/-}$ ES cells. Expression levels (mean ± standard error) were normalized to *β-Actin* and expressed relative to wild-type control ES cells. MMERVK10C env.c and LINE1 5'UTR primer sets (Supplementary Figure S7) were used to assess *MMERVK10C* and *LINE-1* expression.The asterisk indicates a statistically significant difference ($p<0.05$). Note that different primers for *RLTR4* elements behave differently in the qRT-PCR assay. (C) qRT-PCR for different *MMERVK10C* primer sets in *Tex19.1*$^{-/-}$ knockout and littermate control testes at 16 dpp. Expression levels (mean ± standard error for three animals) were normalized to *β-Actin* and expressed relative to littermate controls. Asterisks indicate statistically significant differences ($p<0.05$) (D) Plot showing the *MMERVK10C* genomic contigs flanked by *RLTR10C* LTRs that match only upregulated probes (blue), only unaffected probes (brown), neither class of probes (grey), or both classes of probe (green) in *Tex19.1*$^{-/-}$ testes. Each contig is represented by a horizontal line that indicates the regions of the *MMERVK10C* sequence within it. The upregulated *MMERVK10C* contigs appear to contain recurrent deletions and may be non-autonomous. The positions of the qRT-PCR primers used in (C) are shaded orange. (E) Plot showing the bimodal behaviour of *IAP-int* retrotransposon probe populations in Dnmt TKO ES cells. Vertical lines indicate a 4 fold change. (F) qRT-PCR for of repetitive elements in Dnmt TKO ES cells. Expression levels (mean ± standard error) were normalized to *Gapdh* and expressed relative to wild-type control ES cells. The asterisk indicates a statistically significant difference ($p<0.05$). The LINE1 5'UTR.b primer set (Supplementary Figure S7) was used to assess *LINE-1* expression. Note the difference in behaviour between the two IAP-int primer sets. The *IAP* contig carrying deletions in the AP-1 binding site shown in panel G (IAP_chr10 primers) is expressed but not upregulated in

Dnmt TKO ES cells. (G) Sequence alignment between an LTR of a full-length *IAP* element that does not change expression in Dnmt TKO ES cells (IAP_chr10), and the consensus sequence for the LTR (IAPLTR1a_Mm). The 10 bp deletion removes the AP-1 transcription factor binding site in the LTR.

Supporting Information

Supporting Figure S1 – repeat_organization.pdf

pdf showing a schematic diagram of Repeatmasker organization of murine repetitive elements into classes and families

Supporting Dataset S2– Ms6V1_repprobes.txt

Tab-delimited text file containing complementary repeat probes in Illumina Mouse 6V1 Beadchips

Supporting Dataset S3 - Ms6V1_1_repprobes.txt

Tab-delimited text file containing complementary repeat probes in Illumina Mouse 6V1.1 Beadchips

Supporting Dataset S4 - Ms6V2_repprobes.txt

Tab-delimited text file containing complementary repeat probes in Illumina Mouse 6V2 Beadchips

Supporting Dataset S5 - MG_U74Av2_repprobes.txt

Tab-delimited text file containing complementary repeat probes in Affymetrix Murine Genome U74Av2 GeneChips

Supporting Dataset S6 - Mouse430_2_repprobes.txt

Tab-delimited text file containing complementary repeat probes in Affymetrix Mouse430 2.0 GeneChips

Supporting Figure S7- primer_sequences.pdf

pdf showing a the sequences of primers used for qRT-PCR in this study

Tables

	Illumina			Affymetrix	
	WG6 v1.0	WG6 v1.1	WG6 v2.0	U74Av2	430 2.0
All probes	46,005	46,632	45,281	197,993	496,468
Probes matching repetitive elements (non-complementary)	899	912	867	2,636	19,870
Probes matching repetitive elements (complementary)	1,397	1,425	1,438	4,239	26,124

Table 1. Number of Probes Matching Repetitive Elements in Mouse Gene Expression Microarray Platforms

	Mouse Genome	Illumina			Affymetrix	
		WG-6 v1.0	WG-6 v1.1	WG-6 v2.0	U74Av2	430 2.0
LINE	**122 elements** 1.3 million loci	**70 elements** 351 probes	**71 elements** 358 probes	**66 elements** 321 probes	**62 elements** 631 probes	**97 elements** 4635 probes
SINE	**41 elements** 2.1 million loci	**30 elements** 473 probes	**30 elements** 486 probes	**32 elements** 558 probes	**32 elements** 1465 probes	**37 elements** 11650 probes
LTR	**471 elements** 1.2 million loci	**153 elements** 393 probes	**155 elements** 396 probes	**153 elements** 372 probes	**107 elements** 1362 probes	**291 elements** 7293 probes
DNA	**156 elements** 0.2 million loci	**42 elements** 69 probes	**43 elements** 71 probes	**40 elements** 58 probes	**36 elements** 229 probes	**88 elements** 1329 probes
Satellite	**8 elements** 0.01 million loci	**2 elements** 55 probes	**2 elements** 54 probes	**2 elements** 61 probes	**2 elements** 266 probes	**2 elements** 463 probes
Simple	**315 elements** 1.5 million loci	**8 elements** 9 probes	**3 elements** 9 probes	**9 elements** 12 probes	**26 elements** 67 probes	**47 elements** 168 probes
Other	**108 elements** 0.6 million loci	**13 elements** 47 probes	**14 elements** 51 probes	**13 elements** 56 probes	**15 elements** 219 probes	**32 elements** 586 probes
Total	**1,221 elements** 6.9 million loci	**318 elements** 1397 probes	**323 elements** 1425 probes	**315 elements** 1438 probes	**280 elements** 4239 probes	**594 elements** 26124 probes

Table 2. Number of Different Repetitive Elements Represented by Complementary Probes in Mouse Gene Expression Microarray Platforms. Mouse genome data is derived from Repeatmasker annotation of the mm9 assembly of the sequenced genome downloaded from the UCSC genome browser [62].

		ES cells					Testes
		Dnmt TKO	*Eset*^{shRNA}	*Hdac1*$^{-/-}$	*Ring1B*$^{-/-}$	*Ring1B*$^{-/-}$ *Eed*$^{-/-}$	*Tex19.1*$^{-/-}$
ERV1	*MMERGLN*	↑	↑	–	–	–	–
	RLTR1B	↑	–	–	–	–	–
	RLTR4	–	–	–	(↑)	(↑)	–
	MMVL30	–	–	–	–	↑	–
ERVK	*IAP*	↑	↑	⇩	–	(↑)	–
	RLTR44	–	–	–	↑	↑	–
	RLTR45	↑	↑	↑	–	↑	–
	MMERVK10C	–*	↑	–	–	–	↑
	ETnERV3	–	–	↑	–	–	–
ERVL	*MMERVL*	–	↑	–	–	–	–
LINE	*LINE-1*	–	↑	–	–	–	–

Table 3. Summary of Changes in Repetitive Element Expression Detected by Microarray Repeat-Annotation in this Study. Statistically significant upregulation and downregulation of repetitive element expression in mutant ES cell lines or testes is indicated by up and down arrows respectively. Changes that only appear to affect a small number of probes for a repetitive element are indicated in brackets. The degree of change in gene expression detected for these elements is detailed in the main text. *Although changes in *MMERVK10C* expression were not detected in Dnmt TKO ES cell microrray data in this study, RNA-seq analysis suggests that some genomic copies of *MMERVK10C* are upregulated in Dnmt TKO ES cells [22].